(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 January 2002 (10.01.2002)

PCT

(10) International Publication Number
WO 02/02623 A2

(51) International Patent Classification[7]: C07K 14/47

(21) International Application Number: PCT/US01/20975

(22) International Filing Date: 28 June 2001 (28.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/215,696	29 June 2000 (29.06.2000)	US
60/227,142	22 August 2000 (22.08.2000)	US
60/230,481	6 September 2000 (06.09.2000)	US
60/257,729	21 December 2000 (21.12.2000)	US

(71) Applicant *(for all designated States except US)*: **CORIXA CORPORATION** [US/US]; 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **WANG, Tongtong** [US/US]; 8049 NE 28th Street, Seattle, WA 98039 (US). **MCNEILL, Patricia, D.** [US/US]; 1333 South 290th Place, Federal Way, WA 98003 (US). **WANTANABE, Yoshihiro** [JP/US]; 2266 78th Avenue SE, Mercer Island, WA 98040 (US). **CARTER, Darrick** [US/US]; 321 Summit Avenue East, Seattle, WA 98102 (US). **HENDERSON, Robert, A.** [US/US]; 8904 192nd Street SW, Edmonds, WA 98026 (US). **KALOS, Michael, D.** [US/US]; 8116 Dayton Avenue North, Seattle, WA 98103 (US).

(74) Agents: **VERNA, James, M.** et al.; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/02623 A2

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, such as lung cancer, are disclosed. Compositions may comprise one or more lung tumor proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a lung tumor protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as lung cancer. Diagnostic methods based on detecting a lung tumor protein, or mRNA encoding such a protein, in a sample are also provided.

COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER

TECHNICAL FIELD OF THE INVENTION

The present invention relates generally to therapy and diagnosis of cancer, such as lung cancer. The invention is more specifically related to polypeptides comprising at least a portion of a lung tumor protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides may be used in vaccines and pharmaceutical compositions for prevention and treatment of lung cancer and for the diagnosis and monitoring of such cancers.

BACKGROUND OF THE INVENTION

Cancer is a significant health problem throughout the world. Although advances have been made in detection and therapy of cancer, no vaccine or other universally successful method for prevention or treatment is currently available.

Lung cancer is the primary cause of cancer death among both men and women in the U.S. The five-year survival rate among all lung cancer patients, regardless of the stage of disease at diagnosis, is only 13%. This contrasts with a five-year survival rate of 46% among cases detected while the disease is still localized. However, only 16% of lung cancers are discovered before the disease has spread.

Early detection is difficult since clinical symptoms are often not seen until the disease has reached an advanced stage. Currently, diagnosis is aided by the use of chest x-rays, analysis of the type of cells contained in sputum and fiberoptic examination of the bronchial passages. Treatment regimens are determined by the type and stage of the cancer, and include surgery, radiation therapy and/or chemotherapy.

In spite of considerable research into therapies for these and other cancers, lung remains difficult to diagnose and treat effectively. Accordingly, there is a need in the art for improved methods for detecting and treating such cancers. The present invention fulfills these needs and further provides other related advantages.

SUMMARY OF THE INVENTION

Briefly stated, the present invention provides compositions and methods for the diagnosis and therapy of cancer, such as lung cancer. In one aspect, the present invention provides polypeptides comprising at least a portion of a lung tumor protein, or a variant thereof. Certain portions and other variants are immunogenic, such that the ability of the variant to react with antigen-specific antisera is not substantially diminished. Within certain embodiments, the polypeptide comprises an amino acid sequence selected from the group consisting of (a) SEQ ID NOs:452, 454, 457, and 459-473; (b) a sequence that is encoded by a polynucleotide sequence recited in SEQ ID NO: 1-451, 453, 455-456, and 458; (c) variants of a sequence recited in SEQ ID NO: 1-451, 453, 455-456, and 458; and (d) complements of a sequence of (a) or (b).

The present invention further provides polynucleotides that encode a polypeptide as described above, or a portion thereof (such as a portion encoding at least 15 amino acid residues of a lung tumor protein), expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, vaccines for prophylactic or therapeutic use are provided. Such vaccines comprise a polypeptide or polynucleotide as described above and an immunostimulant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a lung tumor protein; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, vaccines are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins.

Within related aspects, pharmaceutical compositions comprising a fusion protein, or a polynucleotide encoding a fusion protein, in combination with a physiologically acceptable carrier are provided.

Vaccines are further provided, within other aspects, that comprise a fusion protein, or a polynucleotide encoding a fusion protein, in combination with an immunostimulant.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition or vaccine as recited above. The patient may be afflicted with lung cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a lung tumor protein, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a lung tumor protein, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation

and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating $CD4^+$ and/or $CD8^+$ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of a lung tumor protein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for determining the presence or absence of a cancer in a patient, comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody. The cancer may be lung cancer.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount

detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a lung tumor protein; (b) detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a lung tumor protein; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description. All references disclosed herein are

hereby incorporated by reference in their entirety as if each was incorporated individually.

SEQUENCE IDENTIFIERS

SEQ ID NO:1 is the determined cDNA sequence for R0119:A02
SEQ ID NO:2 is the determined cDNA sequence for R0119:A06
SEQ ID NO:3 is the determined cDNA sequence for R0119:A09
SEQ ID NO:4 is the determined cDNA sequence for R0119:A10
SEQ ID NO:5 is the determined cDNA sequence for R0119:A12
SEQ ID NO:6 is the determined cDNA sequence for R0119:B02
SEQ ID NO:7 is the determined cDNA sequence for R0119:B04
SEQ ID NO:8 is the determined cDNA sequence for R0119:B10
SEQ ID NO:9 is the determined cDNA sequence for R0119:C12
SEQ ID NO:10 is the determined cDNA sequence for R0119:D02
SEQ ID NO:11 is the determined cDNA sequence for R0119:D06
SEQ ID NO:12 is the determined cDNA sequence for R0119:D09
SEQ ID NO:13 is the determined cDNA sequence for R0119:D11
SEQ ID NO:14 is the determined cDNA sequence for R0119:D12
SEQ ID NO:15 is the determined cDNA sequence for R0119:E02
SEQ ID NO:16 is the determined cDNA sequence for R0119:E04
SEQ ID NO:17 is the determined cDNA sequence for R0119:E05
SEQ ID NO:18 is the determined cDNA sequence for R0119:E12
SEQ ID NO:19 is the determined cDNA sequence for R0119:F01
SEQ ID NO:20 is the determined cDNA sequence for R0119:F07
SEQ ID NO:21 is the determined cDNA sequence for R0119:F08
SEQ ID NO:22 is the determined cDNA sequence for R0119:F09
SEQ ID NO:23 is the determined cDNA sequence for R0119:F10
SEQ ID NO:24 is the determined cDNA sequence for R0119:F11
SEQ ID NO:25 is the determined cDNA sequence for R0119:F12
SEQ ID NO:26 is the determined cDNA sequence for R0119:G07

SEQ ID NO:27 is the determined cDNA sequence for R0119:G10

SEQ ID NO:28 is the determined cDNA sequence for R0119:H09

SEQ ID NO:29 is the determined cDNA sequence for R0120:A02

SEQ ID NO:30 is the determined cDNA sequence for R0120:A05

SEQ ID NO:31 is the determined cDNA sequence for R0120:A06

SEQ ID NO:32 is the determined cDNA sequence for R0120:A09

SEQ ID NO:33 is the determined cDNA sequence for R0120:A10

SEQ ID NO:34 is the determined cDNA sequence for R0120:A12

SEQ ID NO:35 is the determined cDNA sequence for R0120:B02

SEQ ID NO:36 is the determined cDNA sequence for R0120:B07

SEQ ID NO:37 is the determined cDNA sequence for R0120:B08

SEQ ID NO:38 is the determined cDNA sequence for R0120:B10

SEQ ID NO:39 is the determined cDNA sequence for R0120:C03

SEQ ID NO:40 is the determined cDNA sequence for R0120:C06

SEQ ID NO:41 is the determined cDNA sequence for R0120:C12

SEQ ID NO:42 is the determined cDNA sequence for R0120:D01

SEQ ID NO:43 is the determined cDNA sequence for R0120:D02

SEQ ID NO:44 is the determined cDNA sequence for R0120:D03

SEQ ID NO:45 is the determined cDNA sequence for R0120:D05

SEQ ID NO:46 is the determined cDNA sequence for R0120:D06

SEQ ID NO:47 is the determined cDNA sequence for R0120:D07

SEQ ID NO:48 is the determined cDNA sequence for R0120:D11

SEQ ID NO:49 is the determined cDNA sequence for R0120:D12

SEQ ID NO:50 is the determined cDNA sequence for R0120:E05

SEQ ID NO:51 is the determined cDNA sequence for R0120:E07

SEQ ID NO:52 is the determined cDNA sequence for R0120:E12

SEQ ID NO:53 is the determined cDNA sequence for R0120:F02

SEQ ID NO:54 is the determined cDNA sequence for R0120:F04

SEQ ID NO:55 is the determined cDNA sequence for R0120:F07

SEQ ID NO:56 is the determined cDNA sequence for R0120:F11

SEQ ID NO:57 is the determined cDNA sequence for R0120:G01
SEQ ID NO:58 is the determined cDNA sequence for R0120:G08
SEQ ID NO:59 is the determined cDNA sequence for R0120:H09
SEQ ID NO:60 is the determined cDNA sequence for R0120:H10
SEQ ID NO:61 is the determined cDNA sequence for R0121:A02
SEQ ID NO:62 is the determined cDNA sequence for R0121:A11
SEQ ID NO:63 is the determined cDNA sequence for R0121:B01
SEQ ID NO:64 is the determined cDNA sequence for R0121:B03
SEQ ID NO:65 is the determined cDNA sequence for R0121:B04
SEQ ID NO:66 is the determined cDNA sequence for R0121:C05
SEQ ID NO:67 is the determined cDNA sequence for R0121:C06
SEQ ID NO:68 is the determined cDNA sequence for R0121:D02
SEQ ID NO:69 is the determined cDNA sequence for R0121:D11
SEQ ID NO:70 is the determined cDNA sequence for R0121:E05
SEQ ID NO:71 is the determined cDNA sequence for R0121:E09
SEQ ID NO:72 is the determined cDNA sequence for R0121:E12
SEQ ID NO:73 is the determined cDNA sequence for R0121:F02
SEQ ID NO:74 is the determined cDNA sequence for R0121:F07
SEQ ID NO:75 is the determined cDNA sequence for R0121:G03
SEQ ID NO:76 is the determined cDNA sequence for R0121:G05
SEQ ID NO:77 is the determined cDNA sequence for R0121:H02
SEQ ID NO:78 is the determined cDNA sequence for R0121:H05
SEQ ID NO:79 is the determined cDNA sequence for R0121:G08
SEQ ID NO:80 is the determined cDNA sequence for R0122:A03
SEQ ID NO:81 is the determined cDNA sequence for R0122:A06
SEQ ID NO:82 is the determined cDNA sequence for R0122:A09
SEQ ID NO:83 is the determined cDNA sequence for R0122:B02
SEQ ID NO:84 is the determined cDNA sequence for R0122:B06
SEQ ID NO:85 is the determined cDNA sequence for R0122:B09
SEQ ID NO:86 is the determined cDNA sequence for R0122:B10

SEQ ID NO:87 is the determined cDNA sequence for R0122:C02
SEQ ID NO:88 is the determined cDNA sequence for R0122:C03
SEQ ID NO:89 is the determined cDNA sequence for R0122:C05
SEQ ID NO:90 is the determined cDNA sequence for R0122:C07
SEQ ID NO:91 is the determined cDNA sequence for R0122:C10
SEQ ID NO:92 is the determined cDNA sequence for R0122:C11
SEQ ID NO:93 is the determined cDNA sequence for R0122:D05
SEQ ID NO:94 is the determined cDNA sequence for R0122:D06
SEQ ID NO:95 is the determined cDNA sequence for R0122:D07
SEQ ID NO:96 is the determined cDNA sequence for R0122:E03
SEQ ID NO:97 is the determined cDNA sequence for R0122:G02
SEQ ID NO:98 is the determined cDNA sequence for R0122:F03
SEQ ID NO:99 is the determined cDNA sequence for R0122:F05
SEQ ID NO:100 is the determined cDNA sequence for R0122:F07
SEQ ID NO:101 is the determined cDNA sequence for R0122:F08
SEQ ID NO:102 is the determined cDNA sequence for R0122:F09
SEQ ID NO:103 is the determined cDNA sequence for R0122:F10
SEQ ID NO:104 is the determined cDNA sequence for R0122:G05
SEQ ID NO:105 is the determined cDNA sequence for R0122:G06
SEQ ID NO:106 is the determined cDNA sequence for R0122:G08
SEQ ID NO:107 is the determined cDNA sequence for R0122:G09
SEQ ID NO:108 is the determined cDNA sequence for R0122:G10
SEQ ID NO:109 is the determined cDNA sequence for R0122:G11
SEQ ID NO:110 is the determined cDNA sequence for R0122:G12
SEQ ID NO:111 is the determined cDNA sequence for R0122:H02
SEQ ID NO:112 is the determined cDNA sequence for R0122:H03
SEQ ID NO:113 is the determined cDNA sequence for R0122:H06
SEQ ID NO:114 is the determined cDNA sequence for R0122:H07
SEQ ID NO:115 is the determined cDNA sequence for R0122:H08
SEQ ID NO:116 is the determined cDNA sequence for R0122:H09

SEQ ID NO:117 is the determined cDNA sequence for R0123:A02
SEQ ID NO:118 is the determined cDNA sequence for R0123:A09
SEQ ID NO:119 is the determined cDNA sequence for R0123:B03
SEQ ID NO:120 is the determined cDNA sequence for R0123:B04
SEQ ID NO:121 is the determined cDNA sequence for R0123:B07
SEQ ID NO:122 is the determined cDNA sequence for R0123:B08
SEQ ID NO:123 is the determined cDNA sequence for R0123:C03
SEQ ID NO:124 is the determined cDNA sequence for R0123:C04
SEQ ID NO:125 is the determined cDNA sequence for R0123:C07
SEQ ID NO:126 is the determined cDNA sequence for R0123:D03
SEQ ID NO:127 is the determined cDNA sequence for R0123:D05
SEQ ID NO:128 is the determined cDNA sequence for R0123:D07
SEQ ID NO:129 is the determined cDNA sequence for R0123:D09
SEQ ID NO:130 is the determined cDNA sequence for R0123:D10
SEQ ID NO:131 is the determined cDNA sequence for R0123:E04
SEQ ID NO:132 is the determined cDNA sequence for R0123:F01
SEQ ID NO:133 is the determined cDNA sequence for R0123:F03
SEQ ID NO:134 is the determined cDNA sequence for R0123:F04
SEQ ID NO:135 is the determined cDNA sequence for R0123:F10
SEQ ID NO:136 is the determined cDNA sequence for R0123:G03
SEQ ID NO:137 is the determined cDNA sequence for R0123:G11
SEQ ID NO:138 is the determined cDNA sequence for R0123:H04
SEQ ID NO:139 is the determined cDNA sequence for R0123:H05
SEQ ID NO:140 is the determined cDNA sequence for R0123:H08
SEQ ID NO:141 is the determined cDNA sequence for R0123:H09
SEQ ID NO:142 is the determined cDNA sequence for R0123:HH11
SEQ ID NO:143 is the determined cDNA sequence for R0124:A06
SEQ ID NO:144 is the determined cDNA sequence for R0124:A07
SEQ ID NO:145 is the determined cDNA sequence for R0124:A09
SEQ ID NO:146 is the determined cDNA sequence for R0124:B02

SEQ ID NO:147 is the determined cDNA sequence for R0124:B06
SEQ ID NO:148 is the determined cDNA sequence for R0124:B07
SEQ ID NO:149 is the determined cDNA sequence for R0124:B08
SEQ ID NO:150 is the determined cDNA sequence for R0124:C02
SEQ ID NO:151 is the determined cDNA sequence for R0124:C04
SEQ ID NO:152 is the determined cDNA sequence for R0124:C06
SEQ ID NO:153 is the determined cDNA sequence for R0124:C07
SEQ ID NO:154 is the determined cDNA sequence for R0124:D02
SEQ ID NO:155 is the determined cDNA sequence for R0124:D10
SEQ ID NO:156 is the determined cDNA sequence for R0124:E03
SEQ ID NO:157 is the determined cDNA sequence for R0159:A02
SEQ ID NO:158 is the determined cDNA sequence for R0159:A03
SEQ ID NO:159 is the determined cDNA sequence for R0159:A06
SEQ ID NO:160 is the determined cDNA sequence for R0159:A07
SEQ ID NO:161 is the determined cDNA sequence for R0159:A09
SEQ ID NO:162 is the determined cDNA sequence for R0159:A10
SEQ ID NO:163 is the determined cDNA sequence for R0159:A11
SEQ ID NO:164 is the determined cDNA sequence for R0159:A12
SEQ ID NO:165 is the determined cDNA sequence for R0159:B01
SEQ ID NO:166 is the determined cDNA sequence for R0159:B02
SEQ ID NO:167 is the determined cDNA sequence for R0159:B03
SEQ ID NO:168 is the determined cDNA sequence for R0159:B04
SEQ ID NO:169 is the determined cDNA sequence for R0159:B05
SEQ ID NO:170 is the determined cDNA sequence for R0159:B08
SEQ ID NO:171 is the determined cDNA sequence for R0159:B11
SEQ ID NO:172 is the determined cDNA sequence for R0159:C02
SEQ ID NO:173 is the determined cDNA sequence for R0159:C05
SEQ ID NO:174 is the determined cDNA sequence for R0159:C09
SEQ ID NO:175 is the determined cDNA sequence for R0159:C10
SEQ ID NO:176 is the determined cDNA sequence for R0159:D04

SEQ ID NO:177 is the determined cDNA sequence for R0159:D09
SEQ ID NO:178 is the determined cDNA sequence for R0159:D10
SEQ ID NO:179 is the determined cDNA sequence for R0159:D11
SEQ ID NO:180 is the determined cDNA sequence for R0159:E05
SEQ ID NO:181 is the determined cDNA sequence for R0159:E08
SEQ ID NO:182 is the determined cDNA sequence for R0159:F03
SEQ ID NO:183 is the determined cDNA sequence for R0159:F08
SEQ ID NO:184 is the determined cDNA sequence for R0159:F10
SEQ ID NO:185 is the determined cDNA sequence for R0159:F11
SEQ ID NO:186 is the determined cDNA sequence for R0159:F12
SEQ ID NO:187 is the determined cDNA sequence for R0159:G01
SEQ ID NO:188 is the determined cDNA sequence for R0159:G03
SEQ ID NO:189 is the determined cDNA sequence for R0159:G06
SEQ ID NO:190 is the determined cDNA sequence for R0159:G08
SEQ ID NO:191 is the determined cDNA sequence for R0159:G09
SEQ ID NO:192 is the determined cDNA sequence for R0159:G10
SEQ ID NO:193 is the determined cDNA sequence for R0159:G12
SEQ ID NO:194 is the determined cDNA sequence for R0159:H01
SEQ ID NO:195 is the determined cDNA sequence for R0159:H02
SEQ ID NO:196 is the determined cDNA sequence for R0159:H07
SEQ ID NO:197 is the determined cDNA sequence for R0159:H08
SEQ ID NO:198 is the determined cDNA sequence for R0160:A02
SEQ ID NO:199 is the determined cDNA sequence for R0160:A03
SEQ ID NO:200 is the determined cDNA sequence for R0160:A09
SEQ ID NO:201 is the determined cDNA sequence for R0160:B03
SEQ ID NO:202 is the determined cDNA sequence for R0160:B05
SEQ ID NO:203 is the determined cDNA sequence for R0160:B06
SEQ ID NO:204 is the determined cDNA sequence for R0160:B10
SEQ ID NO:205 is the determined cDNA sequence for R0160:C01
SEQ ID NO:206 is the determined cDNA sequence for R0160:C02

SEQ ID NO:207 is the determined cDNA sequence for R0160:C03
SEQ ID NO:208 is the determined cDNA sequence for R0160:C06
SEQ ID NO:209 is the determined cDNA sequence for R0160:C11
SEQ ID NO:210 is the determined cDNA sequence for R0160:D03
SEQ ID NO:211 is the determined cDNA sequence for R0160:D05
SEQ ID NO:212 is the determined cDNA sequence for R0160:D06
SEQ ID NO:213 is the determined cDNA sequence for R0160:E05
SEQ ID NO:214 is the determined cDNA sequence for R0160:E10
SEQ ID NO:215 is the determined cDNA sequence for R0160:E11
SEQ ID NO:216 is the determined cDNA sequence for R0160:F02
SEQ ID NO:217 is the determined cDNA sequence for R0160:F05
SEQ ID NO:218 is the determined cDNA sequence for R0160:G01
SEQ ID NO:219 is the determined cDNA sequence for R0160:G05
SEQ ID NO:220 is the determined cDNA sequence for R0160:G06
SEQ ID NO:221 is the determined cDNA sequence for R0160:G07
SEQ ID NO:222 is the determined cDNA sequence for R0160:H01
SEQ ID NO:223 is the determined cDNA sequence for R0160:H04
SEQ ID NO:224 is the determined cDNA sequence for R0160:H06
SEQ ID NO:225 is the determined cDNA sequence for R0161:A05
SEQ ID NO:226 is the determined cDNA sequence for R0161:A06
SEQ ID NO:227 is the determined cDNA sequence for R0161:A08
SEQ ID NO:228 is the determined cDNA sequence for R0161:A09
SEQ ID NO:229 is the determined cDNA sequence for R0161:A11
SEQ ID NO:230 is the determined cDNA sequence for R0161:A12
SEQ ID NO:231 is the determined cDNA sequence for R0161:B01
SEQ ID NO:232 is the determined cDNA sequence for R0161:B04
SEQ ID NO:233 is the determined cDNA sequence for R0161:B06
SEQ ID NO:234 is the determined cDNA sequence for R0161:B07
SEQ ID NO:235 is the determined cDNA sequence for R0161:B11
SEQ ID NO:236 is the determined cDNA sequence for R0161:B12

SEQ ID NO:237 is the determined cDNA sequence for R0161:C01
SEQ ID NO:238 is the determined cDNA sequence for R0161:C04
SEQ ID NO:239 is the determined cDNA sequence for R0161:C05
SEQ ID NO:240 is the determined cDNA sequence for R0161:C08
SEQ ID NO:241 is the determined cDNA sequence for R0161:C09
SEQ ID NO:242 is the determined cDNA sequence for R0161:C10
SEQ ID NO:243 is the determined cDNA sequence for R0161:C11
SEQ ID NO:244 is the determined cDNA sequence for R0161:C12
SEQ ID NO:245 is the determined cDNA sequence for R0161:D02
SEQ ID NO:246 is the determined cDNA sequence for R0161:D03
SEQ ID NO:247 is the determined cDNA sequence for R0161:D04
SEQ ID NO:248 is the determined cDNA sequence for R0161:D05
SEQ ID NO:249 is the determined cDNA sequence for R0161:D08
SEQ ID NO:250 is the determined cDNA sequence for R0161:D09
SEQ ID NO:251 is the determined cDNA sequence for R0161:E02
SEQ ID NO:252 is the determined cDNA sequence for R0161:E03
SEQ ID NO:253 is the determined cDNA sequence for R0161:E04
SEQ ID NO:254 is the determined cDNA sequence for R0161:E05
SEQ ID NO:255 is the determined cDNA sequence for R0161:E06
SEQ ID NO:256 is the determined cDNA sequence for R0161:E07
SEQ ID NO:257 is the determined cDNA sequence for R0161:E08
SEQ ID NO:258 is the determined cDNA sequence for R0161:E10
SEQ ID NO:259 is the determined cDNA sequence for R0161:E12
SEQ ID NO:260 is the determined cDNA sequence for R0161:F01
SEQ ID NO:261 is the determined cDNA sequence for R0161:F03
SEQ ID NO:262 is the determined cDNA sequence for R0161:F04
SEQ ID NO:263 is the determined cDNA sequence for R0161:F05
SEQ ID NO:264 is the determined cDNA sequence for R0161:F07
SEQ ID NO:265 is the determined cDNA sequence for R0161:F08
SEQ ID NO:266 is the determined cDNA sequence for R0161:F11

SEQ ID NO:267 is the determined cDNA sequence for R0161:F12
SEQ ID NO:268 is the determined cDNA sequence for R0161:G01
SEQ ID NO:269 is the determined cDNA sequence for R0161:G02
SEQ ID NO:270 is the determined cDNA sequence for R0161:G03
SEQ ID NO:271 is the determined cDNA sequence for R0161:G04
SEQ ID NO:272 is the determined cDNA sequence for R0161:G05
SEQ ID NO:273 is the determined cDNA sequence for R0161:G07
SEQ ID NO:274 is the determined cDNA sequence for R0161:G09
SEQ ID NO:275 is the determined cDNA sequence for R0161:G12
SEQ ID NO:276 is the determined cDNA sequence for R0161:H03
SEQ ID NO:277 is the determined cDNA sequence for R0161:H06
SEQ ID NO:278 is the determined cDNA sequence for R0161:H07
SEQ ID NO:279 is the determined cDNA sequence for R0161:H08
SEQ ID NO:280 is the determined cDNA sequence for R0161:H10
SEQ ID NO:281 is the determined cDNA sequence for R0162:A06
SEQ ID NO:282 is the determined cDNA sequence for R0162:B05
SEQ ID NO:283 is the determined cDNA sequence for R0162:B09
SEQ ID NO:284 is the determined cDNA sequence for R0162:B12
SEQ ID NO:285 is the determined cDNA sequence for R0162:C01
SEQ ID NO:286 is the determined cDNA sequence for R0162:C10
SEQ ID NO:287 is the determined cDNA sequence for R0162:D01
SEQ ID NO:288 is the determined cDNA sequence for R0162:D02
SEQ ID NO:289 is the determined cDNA sequence for R0162:D05
SEQ ID NO:290 is the determined cDNA sequence for R0162:D06
SEQ ID NO:291 is the determined cDNA sequence for R0162:D09
SEQ ID NO:292 is the determined cDNA sequence for R0162:D10
SEQ ID NO:293 is the determined cDNA sequence for R0162:D12
SEQ ID NO:294 is the determined cDNA sequence for R0162:E01
SEQ ID NO:295 is the determined cDNA sequence for R0162:E02
SEQ ID NO:296 is the determined cDNA sequence for R0162:E04

SEQ ID NO:297 is the determined cDNA sequence for R0162:E05
SEQ ID NO:298 is the determined cDNA sequence for R0162:E06
SEQ ID NO:299 is the determined cDNA sequence for R0162:E08
SEQ ID NO:300 is the determined cDNA sequence for R0162:E09
SEQ ID NO:301 is the determined cDNA sequence for R0162:E10
SEQ ID NO:302 is the determined cDNA sequence for R0162:E12
SEQ ID NO:303 is the determined cDNA sequence for R0162:F05
SEQ ID NO:304 is the determined cDNA sequence for R0162:G04
SEQ ID NO:305 is the determined cDNA sequence for R0162:G05
SEQ ID NO:306 is the determined cDNA sequence for R0162:G07
SEQ ID NO:307 is the determined cDNA sequence for R0162:G09
SEQ ID NO:308 is the determined cDNA sequence for R0162:H04
SEQ ID NO:309 is the determined cDNA sequence for R0162:H05
SEQ ID NO:310 is the determined cDNA sequence for R0162:H10
SEQ ID NO:311 is the determined cDNA sequence for R0162:H11
SEQ ID NO:312 is the determined cDNA sequence for R0163:A06
SEQ ID NO:313 is the determined cDNA sequence for R0163:A08
SEQ ID NO:314 is the determined cDNA sequence for R0163:A11
SEQ ID NO:315 is the determined cDNA sequence for R0163:A12
SEQ ID NO:316 is the determined cDNA sequence for R0163:B02
SEQ ID NO:317 is the determined cDNA sequence for R0163:B03
SEQ ID NO:318 is the determined cDNA sequence for R0163:B04
SEQ ID NO:319 is the determined cDNA sequence for R0163:B06
SEQ ID NO:320 is the determined cDNA sequence for R0163:B07
SEQ ID NO:321 is the determined cDNA sequence for R0163:B08
SEQ ID NO:322 is the determined cDNA sequence for R0163:B09
SEQ ID NO:323 is the determined cDNA sequence for R0163:C01
SEQ ID NO:324 is the determined cDNA sequence for R0163:C02
SEQ ID NO:325 is the determined cDNA sequence for R0163:C04
SEQ ID NO:326 is the determined cDNA sequence for R0163:C05

SEQ ID NO:327 is the determined cDNA sequence for R0163:C06
SEQ ID NO:328 is the determined cDNA sequence for R0163:C07
SEQ ID NO:329 is the determined cDNA sequence for R0163:C08
SEQ ID NO:330 is the determined cDNA sequence for R0163:C09
SEQ ID NO:331 is the determined cDNA sequence for R0163:D01
SEQ ID NO:332 is the determined cDNA sequence for R0163:D02
SEQ ID NO:333 is the determined cDNA sequence for R0163:D03
SEQ ID NO:334 is the determined cDNA sequence for R0163:D04
SEQ ID NO:335 is the determined cDNA sequence for R0163:D06
SEQ ID NO:336 is the determined cDNA sequence for R0163:D07
SEQ ID NO:337 is the determined cDNA sequence for R0163:D08
SEQ ID NO:338 is the determined cDNA sequence for R0163:D09
SEQ ID NO:339 is the determined cDNA sequence for R0163:E02
SEQ ID NO:340 is the determined cDNA sequence for R0163:E05
SEQ ID NO:341 is the determined cDNA sequence for R0163:E07
SEQ ID NO:342 is the determined cDNA sequence for R0163:F05
SEQ ID NO:343 is the determined cDNA sequence for R0163:F09
SEQ ID NO:344 is the determined cDNA sequence for R0163:G04
SEQ ID NO:345 is the determined cDNA sequence for R0163:G06
SEQ ID NO:346 is the determined cDNA sequence for R0163:G09
SEQ ID NO:347 is the determined cDNA sequence for R0163:H03
SEQ ID NO:348 is the determined cDNA sequence for R0163:H07
SEQ ID NO:349 is the determined cDNA sequence for R0163:G09
SEQ ID NO:350 is the determined cDNA sequence for R0163:H10
SEQ ID NO:351 is the determined cDNA sequence for R0164:A05
SEQ ID NO:352 is the determined cDNA sequence for R0164:A06
SEQ ID NO:353 is the determined cDNA sequence for R0164:A07
SEQ ID NO:354 is the determined cDNA sequence for R0164:A09
SEQ ID NO:355 is the determined cDNA sequence for R0164:B04
SEQ ID NO:356 is the determined cDNA sequence for R0164:B05

SEQ ID NO:357 is the determined cDNA sequence for R0164:B07
SEQ ID NO:358 is the determined cDNA sequence for R0164:B08
SEQ ID NO:359 is the determined cDNA sequence for R0164:B09
SEQ ID NO:360 is the determined cDNA sequence for R0164:B11
SEQ ID NO:361 is the determined cDNA sequence for R0164:C02
SEQ ID NO:362 is the determined cDNA sequence for R0164:C03
SEQ ID NO:363 is the determined cDNA sequence for R0164:C05
SEQ ID NO:364 is the determined cDNA sequence for R0164:C10
SEQ ID NO:365 is the determined cDNA sequence for R0164:C11
SEQ ID NO:366 is the determined cDNA sequence for R0164:D04
SEQ ID NO:367 is the determined cDNA sequence for R0164:D09
SEQ ID NO:368 is the determined cDNA sequence for R0164:D12
SEQ ID NO:369 is the determined cDNA sequence for R0164:E03
SEQ ID NO:370 is the determined cDNA sequence for R0164:E04
SEQ ID NO:371 is the determined cDNA sequence for R0164:E05
SEQ ID NO:372 is the determined cDNA sequence for R0164:E08
SEQ ID NO:373 is the determined cDNA sequence for R0164:E10
SEQ ID NO:374 is the determined cDNA sequence for R0164:F03
SEQ ID NO:375 is the determined cDNA sequence for R0164:F07
SEQ ID NO:376 is the determined cDNA sequence for R0164:F08
SEQ ID NO:377 is the determined cDNA sequence for R0164:F09
SEQ ID NO:378 is the determined cDNA sequence for R0164:G01
SEQ ID NO:379 is the determined cDNA sequence for R0164:G02
SEQ ID NO:380 is the determined cDNA sequence for R0164:G03
SEQ ID NO:381 is the determined cDNA sequence for R0164:G04
SEQ ID NO:382 is the determined cDNA sequence for R0164:G05
SEQ ID NO:383 is the determined cDNA sequence for R0164:G06
SEQ ID NO:384 is the determined cDNA sequence for R0164:G08
SEQ ID NO:385 is the determined cDNA sequence for R0164:G12
SEQ ID NO:386 is the determined cDNA sequence for R0164:H01

SEQ ID NO:387 is the determined cDNA sequence for R0164:H02
SEQ ID NO:388 is the determined cDNA sequence for R0164:H03
SEQ ID NO:389 is the determined cDNA sequence for R0164:H04
SEQ ID NO:390 is the determined cDNA sequence for R0164:H05
SEQ ID NO:391 is the determined cDNA sequence for R0164:H06
SEQ ID NO:392 is the determined cDNA sequence for R0164:H07
SEQ ID NO:393 is the determined cDNA sequence for R0164:H08
SEQ ID NO:394 is the determined cDNA sequence for R0164:H09
SEQ ID NO:395 is the determined cDNA sequence for R0164:H10
SEQ ID NO:396 is the determined cDNA sequence for R0165:A09
SEQ ID NO:397 is the determined cDNA sequence for R0165:A11
SEQ ID NO:398 is the determined cDNA sequence for R0165:B08
SEQ ID NO:399 is the determined cDNA sequence for R0165:B09
SEQ ID NO:400 is the determined cDNA sequence for R0165:B11
SEQ ID NO:401 is the determined cDNA sequence for R0165:C09
SEQ ID NO:402 is the determined cDNA sequence for R0165:D01
SEQ ID NO:403 is the determined cDNA sequence for R0165:D02
SEQ ID NO:404 is the determined cDNA sequence for R0165:D03
SEQ ID NO:405 is the determined cDNA sequence for R0165:D04
SEQ ID NO:406 is the determined cDNA sequence for R0165:D08
SEQ ID NO:407 is the determined cDNA sequence for R0165:D09
SEQ ID NO:408 is the determined cDNA sequence for R0165:E01
SEQ ID NO:409 is the determined cDNA sequence for R0165:E05
SEQ ID NO:410 is the determined cDNA sequence for R0165:E11
SEQ ID NO:411 is the determined cDNA sequence for R0165:F04
SEQ ID NO:412 is the determined cDNA sequence for R0165:F08
SEQ ID NO:413 is the determined cDNA sequence for R0165:F11
SEQ ID NO:414 is the determined cDNA sequence for R0165:G01
SEQ ID NO:415 is the determined cDNA sequence for R0165:G05
SEQ ID NO:416 is the determined cDNA sequence for R0165:G11

SEQ ID NO:417 is the determined cDNA sequence for R0165:H01
SEQ ID NO:418 is the determined cDNA sequence for R0165:H02
SEQ ID NO:419 is the determined cDNA sequence for R0165:H03
SEQ ID NO:420 is the determined cDNA sequence for R0165:H04
SEQ ID NO:421 is the determined cDNA sequence for R0165:H11
SEQ ID NO:422 is the determined cDNA sequence for '54853.1'
SEQ ID NO:423 is the determined cDNA sequence for '54857.1'
SEQ ID NO:424 is the determined cDNA sequence for '54864.1'
SEQ ID NO:425 is the determined cDNA sequence for '54874.1'
SEQ ID NO:426 is the determined cDNA sequence for '54888.1'
SEQ ID NO:427 is the determined cDNA sequence for '54921.1'
SEQ ID NO:428 is the determined cDNA sequence for '54926.1'
SEQ ID NO:429 is the determined cDNA sequence for '54940.1'
SEQ ID NO:430 is the determined cDNA sequence for '55002.1'
SEQ ID NO:431 is the determined cDNA sequence for '55006.1'
SEQ ID NO:432 is the determined cDNA sequence for '55007.1'
SEQ ID NO:433 is the determined cDNA sequence for '55015.1'
SEQ ID NO:434 is the determined cDNA sequence for '55016.1'
SEQ ID NO:435 is the determined cDNA sequence for '55022.1'
SEQ ID NO:436 is the determined cDNA sequence for '55027.2'
SEQ ID NO:437 is the determined cDNA sequence for '55032.1'
SEQ ID NO:438 is the determined cDNA sequence for '55036.1'
SEQ ID NO:439 is the determined cDNA sequence for '55039.1'
SEQ ID NO:440 is the determined cDNA sequence for 56710.1
SEQ ID NO:441 is the determined cDNA sequence for 56712.1
SEQ ID NO:442 is the determined cDNA sequence for 56716.1
SEQ ID NO:443 is the determined cDNA sequence for 56718.1
SEQ ID NO:444 is the determined cDNA sequence for 56723.1
SEQ ID NO:445 is the determined cDNA sequence for 56724.1
SEQ ID NO:446 is the determined cDNA sequence for 56730.1

SEQ ID NO:447 is the determined cDNA sequence for 56732.1

SEQ ID NO:448 is the determined cDNA sequence for 58375.3

SEQ ID NO:449 is the determined cDNA sequence for 60982.1

SEQ ID NO:450 is the determined cDNA sequence for 60983.2

SEQ ID NO:451 is the determined cDNA sequence for 60983

SEQ ID NO:452 is the amino acid sequence encoded by SEQ ID NO: 451

SEQ ID NO:453 is the determined cDNA sequence for full-length L587S, an extended sequence of clone 55022, SEQ ID NO:435

SEQ ID NO:454 is the amino acid sequence encoded by SEQ ID NO:453

SEQ ID NO:455 is the forward primer PDM-647 for the coding region of clone L587S.

SEQ ID NO:456 is the reverse primer PDM-648 for the coding region of clone L587S.

SEQ ID NO:457 is the amino acid sequence for the expressed recombinant L587S.

SEQ ID NO:458 is the DNA coding sequence for the recombinant L587S.

SEQ ID NO:459 corresponds to amino acids 71-85, an epitope of L587S-specific in the generation of antibodies.

SEQ ID NO:460 corresponds to amino acids 111-125, an epitope of L587S-specific in the generation of antibodies.

SEQ ID NO:461 corresponds to amino acids 1-15, an epitope of L587S-specific in the generation of antibodies.

SEQ ID NO:462 corresponds to amino acids 41-55, an epitope of L587S-specific in the generation of antibodies.

SEQ ID NO:463 corresponds to amino acids 221-235, an epitope of L587S-specific in the generation of antibodies.

SEQ ID NO:464 corresponds to amino acids 171-190, an epitope of L587S-specific in the generation of CD4 T cells.

SEQ ID NO:465 corresponds to amino acids 156-175, an epitope of L587S-specific in the generation of CD4 T cells.

SEQ ID NO:466 corresponds to amino acids 161-180, an epitope of L587S-specific in the generation of CD4 T cells.

SEQ ID NO:467 corresponds to amino acids 166-185, an epitope of L587S-specific in the generation of CD4 T cells.

SEQ ID NO:468 corresponds to amino acids 151-170, an epitope of L587S-specific in the generation of CD4 T cells.

SEQ ID NO:469 corresponds to amino acids 146-165, an epitope of L587S-specific in the generation of CD4 T cells.

SEQ ID NO:470 corresponds to amino acids 41-60, an epitope of L587S-specific in the generation of CD4 T cells.

SEQ ID NO:471 corresponds to amino acids 36-55, an epitope of L587S-specific in the generation of CD4 T cells.

SEQ ID NO:472 corresponds to amino acids 16-35, an epitope of L587S-specific in the generation of CD4 T cells.

SEQ ID NO:473 corresponds to amino acids 11-30, an epitope of L587S-specific in the generation of CD4 T cells.

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions and methods for using the compositions, for example in the therapy and diagnosis of cancer, such as lung cancer. Certain illustrative compositions described herein include lung tumor polypeptides, polynucleotides encoding such polypeptides, binding agents such as antibodies, antigen presenting cells (APCs) and/or immune

system cells (*e.g.*, T cells). A "lung tumor protein," as the term is used herein, refers generally to a protein that is expressed in lung tumor cells at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in a normal tissue, as determined using a representative assay provided herein. Certain lung tumor proteins are tumor proteins that react detectably (within an immunoassay, such as an ELISA or Western blot) with antisera of a patient afflicted with lung cancer.

Therefore, in accordance with the above, and as described further below, the present invention provides illustrative polynucleotide compositions having sequences set forth in SEQ ID NO: 1-451, 453, 455-456, and 458, illustrative polypeptide compositions encoded by the polynucleotide sequences set forth in SEQ ID NO: 1-451, 453, 455-456, and 458 and the amino acid sequences set forth in SEQ ID NO: 452, 454, 457, and 459-473, antibody compositions capable of binding such polypeptides, and numerous additional embodiments employing such compositions, for example in the detection, diagnosis and/or therapy of human lung cancer.

POLYNUCLEOTIDE COMPOSITIONS

As used herein, the terms "DNA segment" and "polynucleotide" refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. Therefore, a DNA segment encoding a polypeptide refers to a DNA segment that contains one or more coding sequences yet is substantially isolated away from, or purified free from, total genomic DNA of the species from which the DNA segment is obtained. Included within the terms "DNA segment" and "polynucleotide" are DNA segments and smaller fragments of such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phagemids, phage, viruses, and the like.

As will be understood by those skilled in the art, the DNA segments of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

"Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA segment does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA segment as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be recognized by the skilled artisan, polynucleotides may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a lung tumor protein or a portion thereof) or may comprise a variant, or a biological or antigenic functional equivalent of such a sequence. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions, as further described below, preferably such that the immunogenicity of the encoded polypeptide is not diminished, relative to a native tumor protein. The effect on the immunogenicity of the encoded polypeptide may generally be assessed as described herein. The term "variants" also encompasses homologous genes of xenogenic origin.

When comparing polynucleotide or polypeptide sequences, two sequences are said to be "identical" if the sequence of nucleotides or amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence

may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenes pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS 5*:151-153; Myers, E.W. and Muller W. (1988) *CABIOS 4*:11-17; Robinson, E.D. (1971) *Comb. Theor 11*:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol. 4*:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA 80*:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.* (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software

for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Therefore, the present invention encompasses polynucleotide and polypeptide sequences having substantial identity to the sequences disclosed herein, for example those comprising at least 50% sequence identity, preferably at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide or polypeptide sequence of this invention using

the methods described herein, (*e.g.*, BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

In additional embodiments, the present invention provides isolated polynucleotides and polypeptides comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative DNA segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

In other embodiments, the present invention is directed to polynucleotides that are capable of hybridizing under moderately stringent conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary

sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS.

Moreover, it will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

PROBES AND PRIMERS

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequence set forth in SEQ ID NO: 1-451 and 453, or to any continuous portion of the sequence, from about 15-25 nucleotides in length up to and including the full length sequence, that one

wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR™ technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered

more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

POLYNUCLEOTIDE IDENTIFICATION AND CHARACTERIZATION

Polynucleotides may be identified, prepared and/or manipulated using any of a variety of well established techniques. For example, a polynucleotide may be identified, as described in more detail below, by screening a microarray of cDNAs for tumor-associated expression (*i.e.*, expression that is at least two fold greater in a tumor than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using a Synteni microarray (Palo Alto, CA) according to the manufacturer's instructions (and essentially as described by Schena *et al.*, *Proc. Natl. Acad. Sci. USA 93*:10614-10619, 1996 and Heller *et al.*, *Proc. Natl. Acad. Sci. USA 94*:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as lung tumor cells. Such polynucleotides may be amplified via polymerase chain reaction (PCR). For this approach, sequence-specific primers may be designed based on the sequences provided herein, and may be purchased or synthesized.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (*e.g.*, a lung tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (*e.g.*, by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or bacteriophage library is then generally screened by hybridizing filters containing

denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (*see* Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, there are numerous amplification techniques for obtaining a full length coding sequence from a partial cDNA sequence. Within such techniques, amplification is generally performed via PCR. Any of a variety of commercially available kits may be used to perform the amplification step. Primers may be designed using, for example, software well known in the art. Primers are preferably 22-30 nucleotides in length, have a GC content of at least 50% and anneal to the target sequence at temperatures of about 68°C to 72°C. The amplified region may be sequenced as described above, and overlapping sequences assembled into a contiguous sequence.

One such amplification technique is inverse PCR (*see* Triglia *et al.*, *Nucl. Acids Res. 16*:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO

96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom *et al.*, *PCR Methods Applic. 1*:111-19, 1991) and walking PCR (Parker *et al.*, *Nucl. Acids. Res. 19*:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (*e.g.*, NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.

POLYNUCLEOTIDE EXPRESSION IN HOST CELLS

In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. *et al.* (1995) *Science 269*:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, T. (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co., New York, N.Y.) or other comparable techniques available in the art. The composition of the synthetic

peptides may be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. *et al.* (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. *et al.* (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York. N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (*e.g.*, baculovirus); plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used.

For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV may be advantageously used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of .beta.-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem. 264*:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel *et al.* (supra) and Grant *et al.* (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used

alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J. 6*:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. *et al.* (1984) *EMBO J. 3*:1671-1680; Broglie, R. *et al.* (1984) *Science 224*:838-843; and Winter, J. *et al.* (1991) *Results Probl. Cell Differ. 17*:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

An insect system may also be used to express a polypeptide of interest. For example, in one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, S. frugiperda cells or Trichoplusia larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. *et al.* (1994) *Proc. Natl. Acad. Sci. 91* :3224-3227).

In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci. 81*:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the

ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. *et al.* (1994) *Results Probl. Cell Differ. 20*:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation. glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. *et al.* (1977) *Cell 11*:223-32) and adenine phosphoribosyltransferase (Lowy, I. *et al.* (1990) *Cell 22*:817-23) genes which can be employed in tk.sup.- or aprt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. *et al.* (1980) *Proc. Natl. Acad. Sci. 77*:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol. 150*:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci. 85*:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. *et al.* (1995) *Methods Mol. Biol. 55*:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include

membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. *et al.* (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul. Minn.) and Maddox, D. E. *et al.* (1983; *J. Exp. Med. 158*:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the

encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen. San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. *et al.* (1992, *Prot. Exp. Purif. 3*:263-281) while the enterokinase cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. *et al.* (1993; *DNA Cell Biol. 12*:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc. 85*:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

SITE-SPECIFIC MUTAGENESIS

Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent polypeptides, through specific

mutagenesis of the underlying polynucleotides that encode them. The technique, well-known to those of skill in the art, further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the antigenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of

a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term "oligonucleotide directed mutagenesis procedure" refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term "oligonucleotide directed mutagenesis procedure" is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment

into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

POLYNUCLEOTIDE AMPLIFICATION TECHNIQUES

A number of template dependent processes are available to amplify the target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCR™) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCR™, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCR™ amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Another method for amplification is the ligase chain reaction (referred to as LCR), disclosed in Eur. Pat. Appl. Publ. No. 320,308 (specifically incorporated herein by reference in its entirety). In LCR, two complementary probe pairs are prepared, and in the presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCR™, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Patent No. 4,883,750, incorporated herein by

reference in its entirety, describes an alternative method of amplification similar to LCR for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880, incorporated herein by reference in its entirety, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA that has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence that can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[α-thio]triphosphates in one strand of a restriction site (Walker *et al.*, 1992, incorporated herein by reference in its entirety), may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, *i.e.* nick translation. A similar method, called Repair Chain Reaction (RCR) is another method of amplification which may be useful in the present invention and is involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA.

Sequences can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having a 3' and 5' sequences of non-target DNA and an internal or "middle" sequence of the target protein specific RNA is hybridized to DNA which is present in a sample. Upon hybridization, the reaction is treated with RNaseH, and the products of the probe are identified as distinctive products by generating a signal that is released after digestion. The original template is annealed to another cycling probe and the reaction is repeated. Thus, CPR involves amplifying a signal generated by hybridization of a probe to a target gene specific expressed nucleic acid.

Still other amplification methods described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025, each of which is incorporated herein by reference in its entirety, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR-like, template and enzyme dependent synthesis. The primers may be modified by labeling with a capture moiety (*e.g.*, biotin) and/or a detector moiety (*e.g.*, enzyme). In the latter application, an excess of labeled probes is added to a sample. In the presence of the target sequence, the probe binds and is cleaved catalytically. After cleavage, the target sequence is released intact to be bound by excess probe. Cleavage of the labeled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (Kwoh *et al.*, 1989; PCT Intl. Pat. Appl. Publ. No. WO 88/10315, incorporated herein by reference in its entirety), including nucleic acid sequence based amplification (NASBA) and 3SR. In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation of a sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques involve annealing a primer that has sequences specific to the target sequence. Following polymerization, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat-denatured again. In either case the single stranded DNA is made fully double stranded by addition of second target-specific primer, followed by polymerization. The double stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNAs are reverse transcribed into DNA, and transcribed once again with a polymerase such as T7 or SP6. The resulting products, whether truncated or complete, indicate target-specific sequences.

Eur. Pat. Appl. Publ. No. 329,822, incorporated herein by reference in its entirety, disclose a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention. The ssRNA is a first template

for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA polymerase). The RNA is then removed from resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for RNA in a duplex with either DNA or RNA). The resultant ssDNA is a second template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to its template. This primer is then extended by DNA polymerase (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting as a double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to make many RNA copies of the DNA. These copies can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

PCT Intl. Pat. Appl. Publ. No. WO 89/06700, incorporated herein by reference in its entirety, disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic; *i.e.* new templates are not produced from the resultant RNA transcripts. Other amplification methods include "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) which are well-known to those of skill in the art.

Methods based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide (Wu and Dean, 1996, incorporated herein by reference in its entirety), may also be used in the amplification of DNA sequences of the present invention.

BIOLOGICAL FUNCTIONAL EQUIVALENTS

Modification and changes may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a polypeptide with desirable characteristics. As mentioned above, it is often desirable to introduce one or more mutations into a specific polynucleotide sequence. In certain circumstances, the resulting encoded polypeptide sequence is altered by this mutation, or in other cases, the sequence of the polypeptide is unchanged by one or more mutations in the encoding polynucleotide.

When it is desirable to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, second-generation molecule, the amino acid changes may be achieved by changing one or more of the codons of the encoding DNA sequence, according to Table 1.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

TABLE 1

Amino Acids			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydropathic index on the basis of its

hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (–0.4); threonine (–0.7); serine (–0.8); tryptophan (–0.9); tyrosine (–1.3); proline (–1.6); histidine (–3.2); glutamate (–3.5); glutamine (–3.5); aspartate (–3.5); asparagine (–3.5); lysine (–3.9); and arginine (–4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

. As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 ± 1); glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (–0.4); proline (–0.5 ± 1); alanine (–0.5); histidine (–0.5); cysteine (–1.0); methionine (–1.3); valine (–1.5); leucine (–1.8); isoleucine (–1.8); tyrosine (–2.3); phenylalanine (–2.5); tryptophan (–3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their

hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl-methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

IN VIVO POLYNUCLEOTIDE DELIVERY TECHNIQUES

In additional embodiments, genetic constructs comprising one or more of the polynucleotides of the invention are introduced into cells *in vivo*. This may be achieved using any of a variety or well known approaches, several of which are outlined below for the purpose of illustration.

1. ADENOVIRUS

One of the preferred methods for *in vivo* delivery of one or more nucleic acid sequences involves the use of an adenovirus expression vector. "Adenovirus expression vector" is meant to include those constructs containing adenovirus sequences sufficient to (a) support packaging of the construct and (b) to express a polynucleotide that has been cloned therein in a sense or antisense orientation. Of course, in the context of an antisense construct, expression does not require that the gene product be synthesized.

The expression vector comprises a genetically engineered form of an adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kb, linear, double-stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kb (Grunhaus and Horwitz, 1992). In contrast to retrovirus,

the adenoviral infection of host cells does not result in chromosomal integration because adenoviral DNA can replicate in an episomal manner without potential genotoxicity. Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification. Adenovirus can infect virtually all epithelial cells regardless of their cell cycle stage. So far, adenoviral infection appears to be linked only to mild disease such as acute respiratory disease in humans.

Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target-cell range and high infectivity. Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are *cis* elements necessary for viral DNA replication and packaging. The early (E) and late (L) regions of the genome contain different transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication, late gene expression and host cell shut-off (Renan, 1990). The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP, (located at 16.8 m.u.) is particularly efficient during the late phase of infection, and all the mRNA's issued from this promoter possess a 5'-tripartite leader (TPL) sequence which makes them preferred mRNA's for translation.

In a current system, recombinant adenovirus is generated from homologous recombination between shuttle vector and provirus vector. Due to the possible recombination between two proviral vectors, wild-type adenovirus may be generated from this process. Therefore, it is critical to isolate a single clone of virus from an individual plaque and examine its genomic structure.

Generation and propagation of the current adenovirus vectors, which are replication deficient, depend on a unique helper cell line, designated 293, which was transformed from human embryonic kidney cells by Ad5 DNA fragments and

constitutively expresses E1 proteins (Graham *et al.*, 1977). Since the E3 region is dispensable from the adenovirus genome (Jones and Shenk, 1978), the current adenovirus vectors, with the help of 293 cells, carry foreign DNA in either the E1, the D3 or both regions (Graham and Prevec, 1991). In nature, adenovirus can package approximately 105% of the wild-type genome (Ghosh-Choudhury *et al.*, 1987), providing capacity for about 2 extra kB of DNA. Combined with the approximately 5.5 kB of DNA that is replaceable in the E1 and E3 regions, the maximum capacity of the current adenovirus vector is under 7.5 kB, or about 15% of the total length of the vector. More than 80% of the adenovirus viral genome remains in the vector backbone and is the source of vector-borne cytotoxicity. Also, the replication deficiency of the E1-deleted virus is incomplete. For example, leakage of viral gene expression has been observed with the currently available vectors at high multiplicities of infection (MOI) (Mulligan, 1993).

Helper cell lines may be derived from human cells such as human embryonic kidney cells, muscle cells, hematopoietic cells or other human embryonic mesenchymal or epithelial cells. Alternatively, the helper cells may be derived from the cells of other mammalian species that are permissive for human adenovirus. Such cells include, *e.g.*, Vero cells or other monkey embryonic mesenchymal or epithelial cells. As stated above, the currently preferred helper cell line is 293.

Recently, Racher *et al.* (1995) disclosed improved methods for culturing 293 cells and propagating adenovirus. In one format, natural cell aggregates are grown by inoculating individual cells into 1 liter siliconized spinner flasks (Techne, Cambridge, UK) containing 100-200 ml of medium. Following stirring at 40 rpm, the cell viability is estimated with trypan blue. In another format, Fibra-Cel microcarriers (Bibby Sterlin, Stone, UK) (5 g/l) is employed as follows. A cell inoculum, resuspended in 5 ml of medium, is added to the carrier (50 ml) in a 250 ml Erlenmeyer flask and left stationary, with occasional agitation, for 1 to 4 h. The medium is then replaced with 50 ml of fresh medium and shaking initiated. For virus production, cells are allowed to grow to about 80% confluence, after which time the medium is replaced (to 25% of the final volume) and adenovirus added at an MOI of 0.05. Cultures are left

stationary overnight, following which the volume is increased to 100% and shaking commenced for another 72 h.

Other than the requirement that the adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not believed to be crucial to the successful practice of the invention. The adenovirus may be of any of the 42 different known serotypes or subgroups A-F. Adenovirus type 5 of subgroup C is the preferred starting material in order to obtain a conditional replication-defective adenovirus vector for use in the present invention, since Adenovirus type 5 is a human adenovirus about which a great deal of biochemical and genetic information is known, and it has historically been used for most constructions employing adenovirus as a vector.

As stated above, the typical vector according to the present invention is replication defective and will not have an adenovirus E1 region. Thus, it will be most convenient to introduce the polynucleotide encoding the gene of interest at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical to the invention. The polynucleotide encoding the gene of interest may also be inserted in lieu of the deleted E3 region in E3 replacement vectors as described by Karlsson *et al.* (1986) or in the E4 region where a helper cell line or helper virus complements the E4 defect.

Adenovirus is easy to grow and manipulate and exhibits broad host range *in vitro* and *in vivo*. This group of viruses can be obtained in high titers, *e.g.*, 10^9-10^{11} plaque-forming units per ml, and they are highly infective. The life cycle of adenovirus does not require integration into the host cell genome. The foreign genes delivered by adenovirus vectors are episomal and, therefore, have low genotoxicity to host cells. No side effects have been reported in studies of vaccination with wild-type adenovirus (Couch *et al.*, 1963; Top *et al.*, 1971), demonstrating their safety and therapeutic potential as *in vivo* gene transfer vectors.

Adenovirus vectors have been used in eukaryotic gene expression (Levrero *et al.*, 1991; Gomez-Foix *et al.*, 1992) and vaccine development (Grunhaus

and Horwitz, 1992; Graham and Prevec, 1992). Recently, animal studies suggested that recombinant adenovirus could be used for gene therapy (Stratford-Perricaudet and Perricaudet, 1991; Stratford-Perricaudet *et al.*, 1990; Rich *et al.*, 1993). Studies in administering recombinant adenovirus to different tissues include trachea instillation (Rosenfeld *et al.*, 1991; Rosenfeld *et al.*, 1992), muscle injection (Ragot *et al.*, 1993), peripheral intravenous injections (Herz and Gerard, 1993) and stereotactic inoculation into the brain (Le Gal La Salle *et al.*, 1993).

2. RETROVIRUSES

The retroviruses are a group of single-stranded RNA viruses characterized by an ability to convert their RNA to double-stranded DNA in infected cells by a process of reverse-transcription (Coffin, 1990). The resulting DNA then stably integrates into cellular chromosomes as a provirus and directs synthesis of viral proteins. The integration results in the retention of the viral gene sequences in the recipient cell and its descendants. The retroviral genome contains three genes, gag, pol, and env that code for capsid proteins, polymerase enzyme, and envelope components, respectively. A sequence found upstream from the gag gene contains a signal for packaging of the genome into virions. Two long terminal repeat (LTR) sequences are present at the 5′ and 3′ ends of the viral genome. These contain strong promoter and enhancer sequences and are also required for integration in the host cell genome (Coffin, 1990).

In order to construct a retroviral vector, a nucleic acid encoding one or more oligonucleotide or polynucleotide sequences of interest is inserted into the viral genome in the place of certain viral sequences to produce a virus that is replication-defective. In order to produce virions, a packaging cell line containing the gag, pol, and env genes but without the LTR and packaging components is constructed (Mann *et al.*, 1983). When a recombinant plasmid containing a cDNA, together with the retroviral LTR and packaging sequences is introduced into this cell line (by calcium phosphate precipitation for example), the packaging sequence allows the RNA transcript of the recombinant plasmid to be packaged into viral particles, which are then secreted into

the culture media (Nicolas and Rubenstein, 1988; Temin, 1986; Mann *et al.*, 1983). The media containing the recombinant retroviruses is then collected, optionally concentrated, and used for gene transfer. Retroviral vectors are able to infect a broad variety of cell types. However, integration and stable expression require the division of host cells (Paskind *et al.*, 1975).

A novel approach designed to allow specific targeting of retrovirus vectors was recently developed based on the chemical modification of a retrovirus by the chemical addition of lactose residues to the viral envelope. This modification could permit the specific infection of hepatocytes *via* sialoglycoprotein receptors.

A different approach to targeting of recombinant retroviruses was designed in which biotinylated antibodies against a retroviral envelope protein and against a specific cell receptor were used. The antibodies were coupled *via* the biotin components by using streptavidin (Roux *et al.*, 1989). Using antibodies against major histocompatibility complex class I and class II antigens, they demonstrated the infection of a variety of human cells that bore those surface antigens with an ecotropic virus *in vitro* (Roux *et al.*, 1989).

3. ADENO-ASSOCIATED VIRUSES

AAV (Ridgeway, 1988; Hermonat and Muzycska, 1984) is a parovirus, discovered as a contamination of adenoviral stocks. It is a ubiquitous virus (antibodies are present in 85% of the US human population) that has not been linked to any disease. It is also classified as a dependovirus, because its replications is dependent on the presence of a helper virus, such as adenovirus. Five serotypes have been isolated, of which AAV-2 is the best characterized. AAV has a single-stranded linear DNA that is encapsidated into capsid proteins VP1, VP2 and VP3 to form an icosahedral virion of 20 to 24 nm in diameter (Muzyczka and McLaughlin, 1988).

The AAV DNA is approximately 4.7 kilobases long. It contains two open reading frames and is flanked by two ITRs (FIG. 2). There are two major genes in the AAV genome: *rep* and *cap*. The *rep* gene codes for proteins responsible for viral replications, whereas *cap* codes for capsid protein VP1-3. Each ITR forms a T-shaped

hairpin structure. These terminal repeats are the only essential *cis* components of the AAV for chromosomal integration. Therefore, the AAV can be used as a vector with all viral coding sequences removed and replaced by the cassette of genes for delivery. Three viral promoters have been identified and named p5, p19, and p40, according to their map position. Transcription from p5 and p19 results in production of rep proteins, and transcription from p40 produces the capsid proteins (Hermonat and Muzyczka, 1984).

There are several factors that prompted researchers to study the possibility of using rAAV as an expression vector. One is that the requirements for delivering a gene to integrate into the host chromosome are surprisingly few. It is necessary to have the 145-bp ITRs, which are only 6% of the AAV genome. This leaves room in the vector to assemble a 4.5-kb DNA insertion. While this carrying capacity may prevent the AAV from delivering large genes, it is amply suited for delivering the antisense constructs of the present invention.

AAV is also a good choice of delivery vehicles due to its safety. There is a relatively complicated rescue mechanism: not only wild type adenovirus but also AAV genes are required to mobilize rAAV. Likewise, AAV is not pathogenic and not associated with any disease. The removal of viral coding sequences minimizes immune reactions to viral gene expression, and therefore, rAAV does not evoke an inflammatory response.

4. OTHER VIRAL VECTORS AS EXPRESSION CONSTRUCTS

Other viral vectors may be employed as expression constructs in the present invention for the delivery of oligonucleotide or polynucleotide sequences to a host cell. Vectors derived from viruses such as vaccinia virus (Ridgeway, 1988; Coupar *et al.*, 1988), lentiviruses, polio viruses and herpes viruses may be employed. They offer several attractive features for various mammalian cells (Friedmann, 1989; Ridgeway, 1988; Coupar *et al.*, 1988; Horwich *et al.*, 1990).

With the recent recognition of defective hepatitis B viruses, new insight was gained into the structure-function relationship of different viral sequences. *In vitro*

studies showed that the virus could retain the ability for helper-dependent packaging and reverse transcription despite the deletion of up to 80% of its genome (Horwich *et al.*, 1990). This suggested that large portions of the genome could be replaced with foreign genetic material. The hepatotropism and persistence (integration) were particularly attractive properties for liver-directed gene transfer. Chang *et al.* (1991) introduced the chloramphenicol acetyltransferase (CAT) gene into duck hepatitis B virus genome in the place of the polymerase, surface, and pre-surface coding sequences. It was cotransfected with wild-type virus into an avian hepatoma cell line. Culture media containing high titers of the recombinant virus were used to infect primary duckling hepatocytes. Stable CAT gene expression was detected for at least 24 days after transfection (Chang *et al.*, 1991).

5. NON-VIRAL VECTORS

In order to effect expression of the oligonucleotide or polynucleotide sequences of the present invention, the expression construct must be delivered into a cell. This delivery may be accomplished *in vitro*, as in laboratory procedures for transforming cells lines, or *in vivo* or *ex vivo*, as in the treatment of certain disease states. As described above, one preferred mechanism for delivery is *via* viral infection where the expression construct is encapsulated in an infectious viral particle.

Once the expression construct has been delivered into the cell the nucleic acid encoding the desired oligonucleotide or polynucleotide sequences may be positioned and expressed at different sites. In certain embodiments, the nucleic acid encoding the construct may be stably integrated into the genome of the cell. This integration may be in the specific location and orientation *via* homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the nucleic acid may be stably maintained in the cell as a separate, episomal segment of DNA. Such nucleic acid segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. How the

expression construct is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of expression construct employed.

In certain embodiments of the invention, the expression construct comprising one or more oligonucleotide or polynucleotide sequences may simply consist of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned above which physically or chemically permeabilize the cell membrane. This is particularly applicable for transfer *in vitro* but it may be applied to *in vivo* use as well. Dubensky *et al.* (1984) successfully injected polyomavirus DNA in the form of calcium phosphate precipitates into liver and spleen of adult and newborn mice demonstrating active viral replication and acute infection. Benvenisty and Reshef (1986) also demonstrated that direct intraperitoneal injection of calcium phosphate-precipitated plasmids results in expression of the transfected genes. It is envisioned that DNA encoding a gene of interest may also be transferred in a similar manner *in vivo* and express the gene product.

Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA-coated microprojectiles to a high velocity allowing them to pierce cell membranes and enter cells without killing them (Klein *et al.*, 1987). Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 1990). The microprojectiles used have consisted of biologically inert substances such as tungsten or gold beads.

Selected organs including the liver, skin, and muscle tissue of rats and mice have been bombarded *in vivo* (Yang *et al.*, 1990; Zelenin *et al.*, 1991). This may require surgical exposure of the tissue or cells, to eliminate any intervening tissue between the gun and the target organ, *i.e. ex vivo* treatment. Again, DNA encoding a particular gene may be delivered *via* this method and still be incorporated by the present invention.

ANTISENSE OLIGONUCLEOTIDES

The end result of the flow of genetic information is the synthesis of protein. DNA is transcribed by polymerases into messenger RNA and translated on the ribosome to yield a folded, functional protein. Thus there are several steps along the route where protein synthesis can be inhibited. The native DNA segment coding for a polypeptide described herein, as all such mammalian DNA strands, has two strands: a sense strand and an antisense strand held together by hydrogen bonding. The messenger RNA coding for polypeptide has the same nucleotide sequence as the sense DNA strand except that the DNA thymidine is replaced by uridine. Thus, synthetic antisense nucleotide sequences will bind to a mRNA and inhibit expression of the protein encoded by that mRNA.

The targeting of antisense oligonucleotides to mRNA is thus one mechanism to shut down protein synthesis, and, consequently, represents a powerful and targeted therapeutic approach. For example, the synthesis of polygalactauronase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829, each specifically incorporated herein by reference in its entirety). Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal $GABA_A$ receptor and human EGF (Jaskulski *et al.*, 1988; Vasanthakumar and Ahmed, 1989; Peris *et al.*, 1998; U. S. Patent 5,801,154; U. S. Patent 5,789,573; U. S. Patent 5,718,709 and U. S. Patent 5,610,288, each specifically incorporated herein by reference in its entirety). Antisense constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683, each specifically incorporated herein by reference in its entirety).

Therefore, in exemplary embodiments, the invention provides oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise

DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein.

Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence (*i.e.* in these illustrative examples the rat and human sequences) and determination of secondary structure, T_m, binding energy, relative stability, and antisense compositions were selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell.

Highly preferred target regions of the mRNA, are those which are at or near the AUG translation initiation codon, and those sequences which were substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations were performed using v.4 of the OLIGO primer analysis software (Rychlik, 1997) and the BLASTN 2.0.5 algorithm software (Altschul *et al.*, 1997).

The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, 1997). It has been demonstrated that several molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane (Morris *et al.*, 1997).

RIBOZYMES

Although proteins traditionally have been used for catalysis of nucleic acids, another class of macromolecules has emerged as useful in this endeavor. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, 1987; Gerlach *et al.*, 1987; Forster and Symons, 1987). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et al.*, 1981; Michel and Westhof, 1990; Reinhold-Hurek and Shub, 1992). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

Ribozyme catalysis has primarily been observed as part of sequence-specific cleavage/ligation reactions involving nucleic acids (Joyce, 1989; Cech *et al.*, 1981). For example, U. S. Patent No. 5,354,855 (specifically incorporated herein by reference) reports that certain ribozymes can act as endonucleases with a sequence specificity greater than that of known ribonucleases and approaching that of the DNA restriction enzymes. Thus, sequence-specific ribozyme-mediated inhibition of gene expression may be particularly suited to therapeutic applications (Scanlon *et al.*, 1991; Sarver *et al.*, 1990). Recently, it was reported that ribozymes elicited genetic changes in some cells lines to which they were applied; the altered genes included the oncogenes H-*ras*, c-*fos* and genes of HIV. Most of this work involved the modification of a target mRNA, based on a specific mutant codon that is cleaved by a specific ribozyme.

Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through

complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its translation) since the concentration of ribozyme necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, 1992). Thus, the specificity of action of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are described by Rossi *et al.* (1992). Examples of hairpin motifs are described by Hampel *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz (1989), Hampel *et al.* (1990) and U. S. Patent 5,631,359 (specifically incorporated herein by reference). An example of the hepatitis δ virus motif is described by Perrotta and Been (1992); an example of the RNaseP motif is described by Guerrier-Takada *et al.* (1983); Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, 1990; Saville and Collins, 1991; Collins and Olive, 1993); and an example of the Group I intron is described in (U. S. Patent 4,987,071, specifically incorporated herein by reference). All that is important in an enzymatic nucleic acid molecule of this invention

is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

In certain embodiments, it may be important to produce enzymatic cleaving agents which exhibit a high degree of specificity for the RNA of a desired target, such as one of the sequences disclosed herein. The enzymatic nucleic acid molecule is preferably targeted to a highly conserved sequence region of a target mRNA. Such enzymatic nucleic acid molecules can be delivered exogenously to specific cells as required. Alternatively, the ribozymes can be expressed from DNA or RNA vectors that are delivered to specific cells.

Small enzymatic nucleic acid motifs (*e.g.*, of the hammerhead or the hairpin structure) may also be used for exogenous delivery. The simple structure of these molecules increases the ability of the enzymatic nucleic acid to invade targeted regions of the mRNA structure. Alternatively, catalytic RNA molecules can be expressed within cells from eukaryotic promoters (*e.g.*, Scanlon *et al.*, 1991; Kashani-Sabet *et al.*, 1992; Dropulic *et al.*, 1992; Weerasinghe *et al.*, 1991; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Sarver *et al.*, 1990). Those skilled in the art realize that any ribozyme can be expressed in eukaryotic cells from the appropriate DNA vector. The activity of such ribozymes can be augmented by their release from the primary transcript by a second ribozyme (Int. Pat. Appl. Publ. No. WO 93/23569, and Int. Pat. Appl. Publ. No. WO 94/02595, both hereby incorporated by reference; Ohkawa *et al.*, 1992; Taira *et al.*, 1991; and Ventura *et al.*, 1993).

Ribozymes may be added directly, or can be complexed with cationic lipids, lipid complexes, packaged within liposomes, or otherwise delivered to target cells. The RNA or RNA complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo* through injection, aerosol inhalation, infusion pump or stent, with or without their incorporation in biopolymers.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Hammerhead or hairpin ribozymes may be individually analyzed by computer folding (Jaeger *et al.*, 1989) to assess whether the ribozyme sequences fold into the appropriate secondary structure. Those ribozymes with unfavorable intramolecular interactions between the binding arms and the catalytic core are eliminated from consideration. Varying binding arm lengths can be chosen to optimize activity. Generally, at least 5 or so bases on each arm are able to bind to, or otherwise interact with, the target RNA.

Ribozymes of the hammerhead or hairpin motif may be designed to anneal to various sites in the mRNA message, and can be chemically synthesized. The method of synthesis used follows the procedure for normal RNA synthesis as described in Usman *et al.* (1987) and in Scaringe *et al.* (1990) and makes use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. Average stepwise coupling yields are typically >98%. Hairpin ribozymes may be synthesized in two parts and annealed to reconstruct an active ribozyme (Chowrira and Burke, 1992). Ribozymes may be modified extensively to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-C-allyl, 2'-flouro, 2'-o-methyl, 2'-H (for a review see *e.g.*, Usman and Cedergren, 1992). Ribozymes may be purified by gel electrophoresis using general methods or by high pressure liquid chromatography and resuspended in water.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Perrault *et al*, 1990; Pieken *et al.*, 1991; Usman and Cedergren, 1992; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur.

Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells (Elroy-Stein and Moss, 1990; Gao and Huang, 1993; Lieber *et al.*, 1993; Zhou *et al.*, 1990).

Ribozymes expressed from such promoters can function in mammalian cells (*e.g.* Kashani-Saber *et al.*, 1992; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Yu *et al.*, 1993; L'Huillier *et al.*, 1992; Lisziewicz *et al.*, 1993). Such transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

Ribozymes may be used as diagnostic tools to examine genetic drift and mutations within diseased cells. They can also be used to assess levels of the target RNA molecule. The close relationship between ribozyme activity and the structure of the target RNA allows the detection of mutations in any region of the molecule which alters the base-pairing and three-dimensional structure of the target RNA. By using multiple ribozymes, one may map nucleotide changes which are important to RNA structure and function *in vitro*, as well as in cells and tissues. Cleavage of target RNAs with ribozymes may be used to inhibit gene expression and define the role (essentially) of specified gene products in the progression of disease. In this manner, other genetic targets may be defined as important mediators of the disease. These studies will lead to better treatment of the disease progression by affording the possibility of combinational therapies (*e.g.*, multiple ribozymes targeted to different genes, ribozymes coupled with known small molecule inhibitors, or intermittent treatment with combinations of ribozymes and/or other chemical or biological molecules). Other *in vitro* uses of ribozymes are well known in the art, and include detection of the presence of mRNA associated with an IL-5 related condition. Such RNA is detected by determining the presence of a cleavage product after treatment with a ribozyme using standard methodology.

PEPTIDE NUCLEIC ACIDS

In certain embodiments, the inventors contemplate the use of peptide nucleic acids (PNAs) in the practice of the methods of the invention. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and

Nielsen, 1997). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (1997) and is incorporated herein by reference. As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter, decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, 1991; Hanvey *et al.*, 1992; Hyrup and Nielsen, 1996; Neilsen, 1996). This chemistry has three important consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc (Dueholm *et al.*, 1994) or Fmoc (Thomson *et al.*, 1995) protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used (Christensen *et al.*, 1995).

PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton *et al.*, 1995). The manual protocol lends itself to the production of chemically modified PNAs or the simultaneous synthesis of families of closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed

by the purification of PNAs by reverse-phase high-pressure liquid chromatography (Norton *et al.*, 1995) providing yields and purity of product similar to those observed during the synthesis of peptides.

Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or for specific functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (Norton *et al.*, 1995; Haaima *et al.*, 1996; Stetsenko *et al.*, 1996; Petersen *et al.*, 1995; Ulmann *et al.*, 1996; Koch *et al.*, 1995; Orum *et al.*, 1995; Footer *et al.*, 1996; Griffith *et al.*, 1995; Kremsky *et al.*, 1996; Pardridge *et al.*, 1995; Boffa *et al.*, 1995; Landsdorp *et al.*, 1996; Gambacorti-Passerini *et al.*, 1996; Armitage *et al.*, 1997; Seeger *et al.*, 1997; Ruskowski *et al.*, 1997). U.S. Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

In contrast to DNA and RNA, which contain negatively charged linkages, the PNA backbone is neutral. In spite of this dramatic alteration, PNAs recognize complementary DNA and RNA by Watson-Crick pairing (Egholm *et al.*, 1993), validating the initial modeling by Nielsen *et al.* (1991). PNAs lack 3′ to 5′ polarity and can bind in either parallel or antiparallel fashion, with the antiparallel mode being preferred (Egholm *et al.*, 1993).

Hybridization of DNA oligonucleotides to DNA and RNA is destabilized by electrostatic repulsion between the negatively charged phosphate backbones of the complementary strands. By contrast, the absence of charge repulsion in PNA-DNA or PNA-RNA duplexes increases the melting temperature (T_m) and reduces the dependence of T_m on the concentration of mono- or divalent cations (Nielsen *et al.*, 1991). The enhanced rate and affinity of hybridization are significant

because they are responsible for the surprising ability of PNAs to perform strand invasion of complementary sequences within relaxed double-stranded DNA. In addition, the efficient hybridization at inverted repeats suggests that PNAs can recognize secondary structure effectively within double-stranded DNA. Enhanced recognition also occurs with PNAs immobilized on surfaces, and Wang *et al.* have shown that support-bound PNAs can be used to detect hybridization events (Wang *et al.*, 1996).

One might expect that tight binding of PNAs to complementary sequences would also increase binding to similar (but not identical) sequences, reducing the sequence specificity of PNA recognition. As with DNA hybridization, however, selective recognition can be achieved by balancing oligomer length and incubation temperature. Moreover, selective hybridization of PNAs is encouraged by PNA-DNA hybridization being less tolerant of base mismatches than DNA-DNA hybridization. For example, a single mismatch within a 16 bp PNA-DNA duplex can reduce the T_m by up to 15°C (Egholm *et al.*, 1993). This high level of discrimination has allowed the development of several PNA-based strategies for the analysis of point mutations (Wang *et al.*, 1996; Carlsson *et al.*, 1996; Thiede *et al.*, 1996; Webb and Hurskainen, 1996; Perry-O'Keefe *et al.*, 1996).

High-affinity binding provides clear advantages for molecular recognition and the development of new applications for PNAs. For example, 11-13 nucleotide PNAs inhibit the activity of telomerase, a ribonucleo-protein that extends telomere ends using an essential RNA template, while the analogous DNA oligomers do not (Norton *et al.*, 1996).

Neutral PNAs are more hydrophobic than analogous DNA oligomers, and this can lead to difficulty solubilizing them at neutral pH, especially if the PNAs have a high purine content or if they have the potential to form secondary structures. Their solubility can be enhanced by attaching one or more positive charges to the PNA termini (Nielsen *et al.*, 1991).

Findings by Allfrey and colleagues suggest that strand invasion will occur spontaneously at sequences within chromosomal DNA (Boffa *et al.*, 1995; Boffa

et al., 1996). These studies targeted PNAs to triplet repeats of the nucleotides CAG and used this recognition to purify transcriptionally active DNA (Boffa *et al.*, 1995) and to inhibit transcription (Boffa *et al.*, 1996). This result suggests that if PNAs can be delivered within cells then they will have the potential to be general sequence-specific regulators of gene expression. Studies and reviews concerning the use of PNAs as antisense and anti-gene agents include Nielsen *et al.* (1993b), Hanvey *et al.* (1992), and Good and Nielsen (1997). Koppelhus *et al.* (1997) have used PNAs to inhibit HIV-1 inverse transcription, showing that PNAs may be used for antiviral therapies.

Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (1993) and Jensen *et al.* (1997). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen *et al.* using BIAcore™ technology.

Other applications of PNAs include use in DNA strand invasion (Nielsen *et al.*, 1991), antisense inhibition (Hanvey *et al.*, 1992), mutational analysis (Orum *et al.*, 1993), enhancers of transcription (Mollegaard *et al.*, 1994), nucleic acid purification (Orum *et al.*, 1995), isolation of transcriptionally active genes (Boffa *et al.*, 1995), blocking of transcription factor binding (Vickers *et al.*, 1995), genome cleavage (Veselkov *et al.*, 1996), biosensors (Wang *et al.*, 1996), *in situ* hybridization (Thisted *et al.*, 1996), and in a alternative to Southern blotting (Perry-O'Keefe, 1996).

POLYPEPTIDE COMPOSITIONS

The present invention, in other aspects, provides polypeptide compositions. Generally, a polypeptide of the invention will be an isolated polypeptide (or an epitope, variant, or active fragment thereof) derived from a mammalian species. Preferably, the polypeptide is encoded by a polynucleotide sequence disclosed herein or a sequence which hybridizes under moderately stringent conditions to a polynucleotide sequence disclosed herein. Alternatively, the polypeptide may be defined as a polypeptide which comprises a contiguous amino acid sequence from an amino acid

sequence disclosed herein, or which polypeptide comprises an entire amino acid sequence disclosed herein.

In the present invention, a polypeptide composition is also understood to comprise one or more polypeptides that are immunologically reactive with antibodies generated against a polypeptide of the invention, particularly a polypeptide encoded by a polynucleotide sequence disclosed in SEQ ID NO: 1-451, 453, 455-456, and 458 or to active fragments, or to variants or biological functional equivalents thereof.

Likewise, a polypeptide composition of the present invention is understood to comprise one or more polypeptides that are capable of eliciting antibodies that are immunologically reactive with one or more polypeptides encoded by one or more contiguous nucleic acid sequences contained in SEQ ID NO: 1-451, 453, 455-456, and 458 or to active fragments, or to variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency.

As used herein, an active fragment of a polypeptide includes a whole or a portion of a polypeptide which is modified by conventional techniques, *e.g.*, mutagenesis, or by addition, deletion, or substitution, but which active fragment exhibits substantially the same structure function, antigenicity, etc., as a polypeptide as described herein.

In certain illustrative embodiments, the polypeptides of the invention will comprise at least an immunogenic portion of a lung tumor protein or a variant thereof, as described herein. As noted above, a "lung tumor protein" is a protein that is expressed by lung tumor cells. Proteins that are lung tumor proteins also react detectably within an immunoassay (such as an ELISA) with antisera from a patient with lung cancer. Polypeptides as described herein may be of any length. Additional sequences derived from the native protein and/or heterologous sequences may be present, and such sequences may (but need not) possess further immunogenic or antigenic properties.

An "immunogenic portion," as used herein is a portion of a protein that is recognized (*i.e.*, specifically bound) by a B-cell and/or T-cell surface antigen

receptor. Such immunogenic portions generally comprise at least 5 amino acid residues, more preferably at least 10, and still more preferably at least 20 amino acid residues of a lung tumor protein or a variant thereof. Certain preferred immunogenic portions include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other preferred immunogenic portions may contain a small N- and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well known techniques. An immunogenic portion of a native lung tumor protein is a portion that reacts with such antisera and/or T-cells at a level that is not substantially less than the reactivity of the full length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Such immunogenic portions may react within such assays at a level that is similar to or greater than the reactivity of the full length polypeptide. Such screens may generally be performed using methods well known to those of ordinary skill in the art, such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Laboratory, 1988. For example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ^{125}I-labeled Protein A.

As noted above, a composition may comprise a variant of a native lung tumor protein. A polypeptide "variant," as used herein, is a polypeptide that differs from a native lung tumor protein in one or more substitutions, deletions, additions and/or insertions, such that the immunogenicity of the polypeptide is not substantially

diminished. In other words, the ability of a variant to react with antigen-specific antisera may be enhanced or unchanged, relative to the native protein, or may be diminished by less than 50%, and preferably less than 20%, relative to the native protein. Such variants may generally be identified by modifying one of the above polypeptide sequences and evaluating the reactivity of the modified polypeptide with antigen-specific antibodies or antisera as described herein. Preferred variants include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other preferred variants include variants in which a small portion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

Polypeptide variants encompassed by the present invention include those exhibiting at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described above) to the polypeptides disclosed herein.

Preferably, a variant contains conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer.

Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (*e.g.*, poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

Polypeptides may be prepared using any of a variety of well known techniques. Recombinant polypeptides encoded by DNA sequences as described above may be readily prepared from the DNA sequences using any of a variety of expression vectors known to those of ordinary skill in the art. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast, and higher eukaryotic cells, such as mammalian cells and plant cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. Supernatants from suitable host/vector systems which secrete recombinant protein or polypeptide into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity matrix or an ion exchange resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

Portions and other variants having less than about 100 amino acids, and generally less than about 50 amino acids, may also be generated by synthetic means, using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. *See* Merrifield, *J. Am. Chem. Soc.* *85*:2149-2146, 1963. Equipment for automated synthesis of polypeptides is

commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

Within certain specific embodiments, a polypeptide may be a fusion protein that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the protein or to enable the protein to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the protein.

Fusion proteins may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion protein is expressed as a recombinant protein, allowing the production of increased levels, relative to a non-fused protein, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion protein that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a

secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea *et al.*, *Gene 40*:39-46, 1985; Murphy *et al.*, *Proc. Natl. Acad. Sci. USA 83*:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

Fusion proteins are also provided. Such proteins comprise a polypeptide as described herein together with an unrelated immunogenic protein. Preferably the immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (*see*, for example, Stoute *et al. New Engl. J. Med.*, *336*:86-91, 1997).

Within preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium Haemophilus influenza B (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (*e.g.*, the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer).

The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemaglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the LytA gene; *Gene 43*:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (*see Biotechnology 10*:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion protein. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

In general, polypeptides (including fusion proteins) and polynucleotides as described herein are isolated. An "isolated" polypeptide or polynucleotide is one that is removed from its original environment. For example, a naturally-occurring protein is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of the natural environment.

BINDING AGENTS

The present invention further provides agents, such as antibodies and antigen-binding fragments thereof, that specifically bind to a lung tumor protein. As

used herein, an antibody, or antigen-binding fragment thereof, is said to "specifically bind" to a lung tumor protein if it reacts at a detectable level (within, for example, an ELISA) with a lung tumor protein, and does not react detectably with unrelated proteins under similar conditions. As used herein, "binding" refers to a noncovalent association between two separate molecules such that a complex is formed. The ability to bind may be evaluated by, for example, determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component concentrations. In general, two compounds are said to "bind," in the context of the present invention, when the binding constant for complex formation exceeds about 10^3 L/mol. The binding constant may be determined using methods well known in the art.

Binding agents may be further capable of differentiating between patients with and without a cancer, such as lung cancer, using the representative assays provided herein. In other words, antibodies or other binding agents that bind to a lung tumor protein will generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. It will be apparent that a statistically significant number of samples with and without the disease should be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.,* Harlow

and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol. 6*:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture

supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

Within certain embodiments, the use of antigen-binding fragments of antibodies may be preferred. Such fragments include Fab fragments, which may be prepared using standard techniques. Briefly, immunoglobulins may be purified from rabbit serum by affinity chromatography on Protein A bead columns (Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988) and digested by papain to yield Fab and Fc fragments. The Fab and Fc fragments may be separated by affinity chromatography on protein A bead columns.

Monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y, ^{123}I, ^{125}I, ^{131}I, ^{186}Re, ^{188}Re, ^{211}At, and ^{212}Bi. Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diptheria toxin, cholera toxin, gelonin, Pseudomonas exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-

containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter *et al.*), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn *et al.*), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler *et al.*).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be

coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato *et al.*), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih *et al.*). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison *et al.* discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous, intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density on the tumor, and the rate of clearance of the antibody.

T CELLS

Immunotherapeutic compositions may also, or alternatively, comprise T cells specific for a lung tumor protein. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO

92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a lung tumor polypeptide, polynucleotide encoding a lung tumor polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide. Preferably, a lung tumor polypeptide or polynucleotide is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a lung tumor polypeptide if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen *et al.*, *Cancer Res.* *54*:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a lung tumor polypeptide (100 ng/ml - 100 μg/ml, preferably 200 ng/ml - 25 μg/ml) for 3 - 7 days should result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN-γ) is indicative of T cell activation (*see* Coligan *et al.*, Current Protocols in Immunology, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a lung tumor polypeptide, polynucleotide or polypeptide-expressing APC may be $CD4^+$ and/or $CD8^+$. Lung tumor protein-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are

derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, $CD4^+$ or $CD8^+$ T cells that proliferate in response to a lung tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a lung tumor polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a lung tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of a lung tumor protein can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

PHARMACEUTICAL COMPOSITIONS

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions disclosed herein in pharmaceutically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will also be understood that, if desired, the nucleic acid segment, RNA, DNA or PNA compositions that express a polypeptide as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA compositions.

Formulation of pharmaceutically-acceptable excipients and carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation.

1. ORAL DELIVERY

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (Mathiowitz *et al.*, 1997; Hwang *et al.*, 1998; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451, each specifically incorporated herein by reference in its entirety). The tablets, troches, pills, capsules and the like may also contain the following: a binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. A syrup of elixir may contain the active compound sucrose as a sweetening agent methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange flavor. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In

addition, the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations may contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared is such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. For example, a mouthwash may be prepared incorporating the active ingredient in the required amount in an appropriate solvent, such as a sodium borate solution (Dobell's Solution). Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

2. INJECTABLE DELIVERY

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even intraperitoneally as described in U. S. Patent 5,543,158; U. S. Patent 5,641,515 and U. S. Patent 5,399,363 (each specifically incorporated herein by reference in its entirety). Solutions of the active compounds as free base or pharmacologically acceptable salts

may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (U. S. Patent 5,466,468, specifically incorporated herein by reference in its entirety). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml

of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject. Moreover, for human administration, preparations should meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

The compositions disclosed herein may be formulated in a neutral or salt form. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug-release capsules, and the like.

As used herein, "carrier" includes any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption

delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified.

3. NASAL DELIVERY

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs *via* nasal aerosol sprays has been described *e.g.*, in U. S. Patent 5,756,353 and U. S. Patent 5,804,212 (each specifically incorporated herein by reference in its entirety). Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, 1998) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871, specifically incorporated herein by reference in its entirety) are also well-known in the pharmaceutical arts. Likewise, transmucosal drug delivery in the form of a polytetrafluoroetheylene support matrix is described in U. S. Patent 5,780,045 (specifically incorporated herein by reference in its entirety).

4. LIPOSOME-, NANOCAPSULE-, AND MICROPARTICLE-MEDIATED DELIVERY

In certain embodiments, the inventors contemplate the use of liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, and the like, for the introduction of the compositions of the present invention into suitable host cells. In

particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like.

Such formulations may be preferred for the introduction of pharmaceutically-acceptable formulations of the nucleic acids or constructs disclosed herein. The formation and use of liposomes is generally known to those of skill in the art (see for example, Couvreur *et al.*, 1977; Couvreur, 1988; Lasic, 1998; which describes the use of liposomes and nanocapsules in the targeted antibiotic therapy for intracellular bacterial infections and diseases). Recently, liposomes were developed with improved serum stability and circulation half-times (Gabizon and Papahadjopoulos, 1988; Allen and Choun, 1987; U. S. Patent 5,741,516, specifically incorporated herein by reference in its entirety). Further, various methods of liposome and liposome like preparations as potential drug carriers have been reviewed (Takakura, 1998; Chandran *et al.*, 1997; Margalit, 1995; U. S. Patent 5,567,434; U. S. Patent 5,552,157; U. S. Patent 5,565,213; U. S. Patent 5,738,868 and U. S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally resistant to transfection by other procedures including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, 1990; Muller *et al.*, 1990). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, drugs (Heath and Martin, 1986; Heath *et al.*, 1986; Balazsovits *et al.*, 1989; Fresta and Puglisi, 1996), radiotherapeutic agents (Pikul *et al.*, 1987), enzymes (Imaizumi *et al.*, 1990a; Imaizumi *et al.*, 1990b), viruses (Faller and Baltimore, 1984), transcription factors and allosteric effectors (Nicolau and Gersonde, 1979) into a variety of cultured cell lines and animals. In addition, several successful clinical trails examining the effectiveness of liposome-mediated drug delivery have been completed (Lopez-Berestein *et al.*, 1985a; 1985b; Coune, 1988; Sculier *et al.*, 1988). Furthermore, several studies suggest that the use of liposomes is not associated with autoimmune responses, toxicity or gonadal localization after systemic delivery (Mori and Fukatsu, 1992).

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs). MLVs generally have diameters of from 25 nm to 4 μm. Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500 Å, containing an aqueous solution in the core.

Liposomes bear resemblance to cellular membranes and are contemplated for use in connection with the present invention as carriers for the peptide compositions. They are widely suitable as both water- and lipid-soluble substances can be entrapped, *i.e.* in the aqueous spaces and within the bilayer itself, respectively. It is possible that the drug-bearing liposomes may even be employed for site-specific delivery of active agents by selectively modifying the liposomal formulation.

In addition to the teachings of Couvreur *et al.* (1977; 1988), the following information may be utilized in generating liposomal formulations. Phospholipids can form a variety of structures other than liposomes when dispersed in water, depending on the molar ratio of lipid to water. At low ratios the liposome is the preferred structure. The physical characteristics of liposomes depend on pH, ionic strength and the presence of divalent cations. Liposomes can show low permeability to ionic and polar substances, but at elevated temperatures undergo a phase transition which markedly alters their permeability. The phase transition involves a change from a closely packed, ordered structure, known as the gel state, to a loosely packed, less-ordered structure, known as the fluid state. This occurs at a characteristic phase-transition temperature and results in an increase in permeability to ions, sugars and drugs.

In addition to temperature, exposure to proteins can alter the permeability of liposomes. Certain soluble proteins, such as cytochrome c, bind, deform and penetrate the bilayer, thereby causing changes in permeability. Cholesterol inhibits this penetration of proteins, apparently by packing the phospholipids more tightly. It is contemplated that the most useful liposome formations for antibiotic and inhibitor delivery will contain cholesterol.

The ability to trap solutes varies between different types of liposomes. For example, MLVs are moderately efficient at trapping solutes, but SUVs are extremely inefficient. SUVs offer the advantage of homogeneity and reproducibility in size distribution, however, and a compromise between size and trapping efficiency is offered by large unilamellar vesicles (LUVs). These are prepared by ether evaporation and are three to four times more efficient at solute entrapment than MLVs.

In addition to liposome characteristics, an important determinant in entrapping compounds is the physicochemical properties of the compound itself. Polar compounds are trapped in the aqueous spaces and nonpolar compounds bind to the lipid bilayer of the vesicle. Polar compounds are released through permeation or when the bilayer is broken, but nonpolar compounds remain affiliated with the bilayer unless it is disrupted by temperature or exposure to lipoproteins. Both types show maximum efflux rates at the phase transition temperature.

Liposomes interact with cells *via* four different mechanisms: endocytosis by phagocytic cells of the reticuloendothelial system such as macrophages and neutrophils; adsorption to the cell surface, either by nonspecific weak hydrophobic or electrostatic forces, or by specific interactions with cell-surface components; fusion with the plasma cell membrane by insertion of the lipid bilayer of the liposome into the plasma membrane, with simultaneous release of liposomal contents into the cytoplasm; and by transfer of liposomal lipids to cellular or subcellular membranes, or vice versa, without any association of the liposome contents. It often is difficult to determine which mechanism is operative and more than one may operate at the same time.

The fate and disposition of intravenously injected liposomes depend on their physical properties, such as size, fluidity, and surface charge. They may persist in tissues for h or days, depending on their composition, and half lives in the blood range from min to several h. Larger liposomes, such as MLVs and LUVs, are taken up rapidly by phagocytic cells of the reticuloendothelial system, but physiology of the circulatory system restrains the exit of such large species at most sites. They can exit only in places where large openings or pores exist in the capillary endothelium, such as the sinusoids of the liver or spleen. Thus, these organs are the predominate site of

uptake. On the other hand, SUVs show a broader tissue distribution but still are sequestered highly in the liver and spleen. In general, this *in vivo* behavior limits the potential targeting of liposomes to only those organs and tissues accessible to their large size. These include the blood, liver, spleen, bone marrow, and lymphoid organs.

Targeting is generally not a limitation in terms of the present invention. However, should specific targeting be desired, methods are available for this to be accomplished. Antibodies may be used to bind to the liposome surface and to direct the antibody and its drug contents to specific antigenic receptors located on a particular cell-type surface. Carbohydrate determinants (glycoprotein or glycolipid cell-surface components that play a role in cell-cell recognition, interaction and adhesion) may also be used as recognition sites as they have potential in directing liposomes to particular cell types. Mostly, it is contemplated that intravenous injection of liposomal preparations would be used, but other routes of administration are also conceivable.

Alternatively, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (Henry-Michelland *et al.*, 1987; Quintanar-Guerrero *et al.*, 1998; Douglas *et al.*, 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention. Such particles may be are easily made, as described (Couvreur *et al.*, 1980; 1988; zur Muhlen *et al.*, 1998; Zambaux *et al.* 1998; Pinto-Alphandry *et al.*, 1995 and U. S. Patent 5,145,684, specifically incorporated herein by reference in its entirety).

VACCINES

In certain preferred embodiments of the present invention, vaccines are provided. The vaccines will generally comprise one or more pharmaceutical compositions, such as those discussed above, in combination with an immunostimulant. An immunostimulant may be any substance that enhances or potentiates an immune

response (antibody and/or cell-mediated) to an exogenous antigen. Examples of immunostimulants include adjuvants, biodegradable microspheres (*e.g.,* polylactic galactide) and liposomes (into which the compound is incorporated; *see e.g.,* Fullerton, U.S. Patent No. 4,235,877). Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Pharmaceutical compositions and vaccines within the scope of the present invention may also contain other compounds, which may be biologically active or inactive. For example, one or more immunogenic portions of other tumor antigens may be present, either incorporated into a fusion polypeptide or as a separate compound, within the composition or vaccine.

Illustrative vaccines may contain DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacteria and viral expression systems. Numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems 15*:143-198, 1998, and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope. In a preferred embodiment, the DNA may be introduced using a viral expression system (*e.g.,* vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for example, in Fisher-Hoch *et al., Proc. Natl. Acad. Sci. USA 86*:317-321, 1989; Flexner *et al., Ann. N.Y. Acad. Sci. 569*:86-103, 1989; Flexner *et al., Vaccine 8*:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques 6*:616-627, 1988; Rosenfeld *et al., Science 252*:431-434, 1991; Kolls *et al., Proc. Natl. Acad. Sci. USA 91*:215-219, 1994; Kass-Eisler *et al., Proc. Natl. Acad.*

Sci. USA 90:11498-11502, 1993; Guzman *et al.*, *Circulation 88*:2838-2848, 1993; and Guzman *et al.*, *Cir. Res. 73*:1202-1207, 1993. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer *et al.*, *Science 259*:1745-1749, 1993 and reviewed by Cohen, *Science 259*:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells. It will be apparent that a vaccine may comprise both a polynucleotide and a polypeptide component. Such vaccines may provide for an enhanced immune response.

It will be apparent that a vaccine may contain pharmaceutically acceptable salts of the polynucleotides and polypeptides provided herein. Such salts may be prepared from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.*, salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (*e.g.*, sodium, potassium, lithium, ammonium, calcium and magnesium salts).

While any suitable carrier known to those of ordinary skill in the art may be employed in the vaccine compositions of this invention, the type of carrier will vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (*e.g.*, polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344 and 5,942,252. One may also employ a carrier comprising the particulate-protein complexes described in U.S. Patent No.

5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

Such compositions may also comprise buffers (*e.g.*, neutral buffered saline or phosphate buffered saline), carbohydrates (*e.g.*, glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (*e.g.*, aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

Any of a variety of immunostimulants may be employed in the vaccines of this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bortadella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

Within the vaccines provided herein, the adjuvant composition is preferably designed to induce an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.*, IFN-γ, TNFα, IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as

provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Corixa Corporation (Seattle, WA; *see* US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato *et al.*, *Science 273*:352, 1996. Another preferred adjuvant is a saponin, preferably QS21 (Aquila Biopharmaceuticals Inc., Framingham, MA), which may be used alone or in combination with other adjuvants. For example, an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Other preferred adjuvants include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (*e.g.*, SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in

pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties.

Any vaccine provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a suitable carrier or excipient. The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule, sponge or gel (composed of polysaccharides, for example) that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology (*see, e.g.*, Coombes *et al.*, *Vaccine 14*:1429-1438, 1996) and administered by, for example, oral, rectal or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane.

Carriers for use within such formulations are biocompatible, and may also be biodegradable; preferably the formulation provides a relatively constant level of active component release. Such carriers include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.*, a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (*see e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and vaccines to facilitate production of an antigen-specific immune response that targets tumor cells. Delivery vehicles include antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve

activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature 392*:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med. 50*:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (*see* Zitvogel *et al.*, *Nature Med. 4*:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNFα to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNFα, CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fcγ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (*e.g.,* CD54 and CD11) and costimulatory molecules (*e.g.,* CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide encoding a lung tumor protein (or portion or other variant thereof) such that the lung tumor polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a composition or vaccine comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi *et al.*, *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the lung tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.,* vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (*e.g.,* a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

Vaccines and pharmaceutical compositions may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are

preferably hermetically sealed to preserve sterility of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a vaccine or pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

CANCER THERAPY

In further aspects of the present invention, the compositions described herein may be used for immunotherapy of cancer, such as lung cancer. Within such methods, pharmaceutical compositions and vaccines are typically administered to a patient. As used herein, a "patient" refers to any warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions and vaccines may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. A cancer may be diagnosed using criteria generally accepted in the art, including the presence of a malignant tumor. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. Administration may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host

immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as $CD8^+$ cytotoxic T lymphocytes and $CD4^+$ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow in vivo and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see*, for example, Cheever *et al.*, *Immunological Reviews 157*:177, 1997).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex*

vivo for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 µg to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a lung tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using

standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

CANCER DETECTION AND DIAGNOSIS

In general, a cancer may be detected in a patient based on the presence of one or more lung tumor proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as lung cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a lung tumor sequence should be present at a level that is at least three fold higher in tumor tissue than in normal tissue

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a

polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length lung tumor proteins and portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 µg, and preferably about 100 ng to about 1 µg, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports

having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (*see, e.g.*, Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with lung cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second

antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as lung cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett *et al.*, *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample

generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1μg, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to

those of ordinary skill in the art that the above protocols may be readily modified to use lung tumor polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such lung tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a lung tumor protein in a biological sample. Within certain methods, a biological sample comprising $CD4^+$ and/or $CD8^+$ T cells isolated from a patient is incubated with a lung tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (*e.g.*, 5 - 25 μg/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of lung tumor polypeptide to serve as a control. For $CD4^+$ T cells, activation is preferably detected by evaluating proliferation of the T cells. For $CD8^+$ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a lung tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a lung tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the lung tumor protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to

polynucleotide encoding a lung tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a lung tumor protein that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence recited in SEQ ID NO: 1-451 and 453. Techniques for both PCR based assays and hybridization assays are well known in the art (*see*, for example, Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.*, *51*:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described

above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple lung tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided herein may be combined with assays for other known tumor antigens.

DIAGNOSTIC KITS

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a lung tumor protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively,

contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a lung tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a lung tumor protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a lung tumor protein.

The following Examples are offered by way of illustration and not by way of limitation.

EXAMPLES

EXAMPLE 1

IDENTIFICATION OF LUNG TUMOR PROTEIN cDNAs

This Example illustrates the identification of cDNA molecules encoding lung tumor proteins.

The cDNAs disclosed herein were generated by sequencing of a subtracted lung squamous tumor cDNA library, LST-S5, and a subtracted metastatic lung adenocarcinoma cDNA library, MS1 (mets3209-S1), as described further below.

TISSUE AND RNA SOURCES

Tumor and some normal tissues used in this studies were from Cooperative Human Tissue Network (CHTN), National Disease Research Interchange (NDRI), and Roswell Park Cancer Center.

CONSTRUCTION OF cDNA LIBRARIES

cDNA libraries were constructed from poly A^+ RNA extracted from a pool of two patient tissues for LST-S5 and a metastatic adenocarcinoma tissue for MS1 using a Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning Kit (GIBCO BRL Life Technologies, Gaithersburg, MD), with modifications. Briefly, BstXI/EcoRI adaptors (Invitrogen, San Diego, CA) were used and cDNA was cloned into pcDNA3.1+ vector (Invitrogen, San Diego, CA) that was digested with BstXI and EcoRI. A total of 1.6×10^6 to 2.7×10^6 independent colonies were obtained for LSCC and lung adenocarcinoma cDNA libraries, with 100% of clones having inserts and the average insert size being 2,100 base pairs.

CONSTRUCTION OF cDNA LIBRARIES USING NORMAL LUNG, HEART AND LIVER TISSUES

Using essentially the same procedure, a normal human lung cDNA library was prepared with a pool of four lung tissue specimens, a normal esophagus cDNA library was prepared from a pool of two esophagus total RNA samples, and a

mixed normal tissue cDNA library was prepared from equal amounts of total RNA isolated from lung, liver, pancreas, skin, brain and PBMC. The normal lung library contained 1.4 x 10^6 independent colonies, with 90% of clones having inserts and the average insert size being 1,800 base pairs. The normal esophagus cDNA library contained 1.0 x 10^6 independent colonies, with 100% of clones having inserts and the average insert size being 1,600 base pairs. The mixed normal tissue cDNA library contained 2.0 x x 10^6 independent colonies, with 100% of clones having inserts and the average insert size being 1,500 base pairs.

LUNG SQUAMOUS CELL CARCINOMA AND LUNG ADENOCARCINOMA-SPECIFIC SUBTRACTED cDNA LIBRARIES

To enrich for genes preferentially expressed in LSCC and/or lung adenocarcinoma, we performed cDNA library subtractions using the above lung squamous cell and adenocarcinoma cDNA libraries as the testers and normal tissue cDNA libraries as driver, as previously described (Sargent and Dawid, 1983; Duguid and Dinauer, 1990), with modifications. Normal lung, esophagus and mixed cDNAs (40μg of each) were digested with BamHI and XhoI, followed by phenol-choloroform extraction and ethanol precipitation. The DNA was then labeled with photoprobe long-arm biotin (Vector Laboratories, Burlingame, CA) and the resulting material was ethanol precipitated and dissolved in H_2O at 2 mg/ml to prepare driver DNA. For tester DNA, 10μg of lung squamous cell carcinoma or lung adenocarcinoma cDNA was digested with NotI and SpeI followed by phenol-chloroform extraction and size fractionation using Chroma spin-400 columns (Clontech, Palo Alto, CA). 5μg tester DNA was mixed with 25μg driver DNA and proceeded for hybridization at 68°C by adding equal volume of 2 X hybridization buffer (1.5M NaCl/10 mM EDTA/50 mM HEPES pH7.5/0.2% sodium dodecyl sulfate). Following hybridization, several rounds of streptavidin treatment and phenol/chloroform extraction were performed to remove biotinlated DNA, both driver DNA and tester DNA hybridizing to driver DNA. The subtracted DNA enriched for tester specific DNA was then hybridized to additional driver DNA for a second round of subtraction. After the second round of subtraction,

DNA was precipitated and ligated into pBCSK+ plasmid vector (Stratagene, La Jolla, CA) to generate a Lung Squamous Tumor-specific Subtracted cDNA library, referred to as LST-5 and a subtracted metastatic lung adenocarcinoma cDNA library, referred to as MS1.

To analyze the subtracted libraries, 20 to 300 clones were randomly picked and plasmid DNA was prepared for sequence analysis with a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A and/or Model 377 (Foster City, CA). These sequences were compared to sequences in the GenBank and human EST databases. The redundancy and the complexity of each subtracted cDNA library was then estimated based on the frequency of each unique cDNA recovered. Highly redundant cDNAs were then used as probes to pre-screen the subtracted cDNA libraries to eliminate redundant cDNA fragments from those to be analyzed by microarray technology.

ANALYSIS OF CDNA EXPRESSION USING MICROARRAY TECHNOLOGY

A total of 672 cDNA sequences isolated in LST-5 and a total of 531 cDNA sequences isolated from MS1 were PCR amplified from individual colonies. Their mRNA expression profiles in lung tumor, normal lung, and other normal and tumor tissues were examined using cDNA microarray technology as described (Shena *et al.*, 1995). In brief, these clones were arrayed onto glass slides as multiple replicas, with each location corresponding to a unique cDNA clone (as many as 5500 clones can be arrayed on a single slide, or chip). Each chip was hybridized with a pair of cDNA probes that were fluorescence-labeled with Cy3 and Cy5, respectively. Typically, 1μg of polyA$^+$ RNA was used to generate each cDNA probe. After hybridization, the chips were scanned and the fluorescence intensity recorded for both Cy3 and Cy5 channels. There were multiple built-in quality control steps. First, the probe quality was monitored using a panel of 18 ubiquitously expressed genes. Secondly, the control plate also had yeast DNA fragments of which complementary RNA was spiked into the probe synthesis for measuring the quality of the probe and the sensitivity of the analysis. Currently, the technology offers a sensitivity of 1 in 100,000 copies of

mRNA. Finally, the reproducibility of this technology was ensured by including duplicated control cDNA elements at different locations. Further validation of the process was indicated in that several differentially expressed genes were identified multiple times in the study, and the expression profiles for these genes are very comparable (not shown).

The following results were obtained and shown in Table 2:

Table 2:

SEQ ID NO:	Ref No:	Element (96)	Ratio	Median Signal 1	Median Signal 2
422	54853	R0120 B7	2.35	0.073	0.031
423	54857	R0120 D1	52.52	4.275	0.081
424	54864	R0120 F4	40.33	5.485	0.136
425	54874	R0120 H4	4.41	0.094	0.021
426	54888	R0121 E12	5.6	0.478	0.085
427	54921	R0123 A11	3.87	0.382	0.099
428	54926	R0123 D5	5.86	0.499	0.085
429	54940	R0123 H11	2.03	0.231	0.114
430	55002	R0124 C11	5.77	0.504	0.087
431	55006	R0124 E3/MS1	2.45	0.182	0.074
432	55007	R0159 E2	2.87	0.473	0.165
433	55015	R0160 B1	8.19	0.451	0.055
434	55016	R0160 C8	2.19	0.165	0.075
435	55022	R0160 G5	3.83	0.121	0.032
436	55027	R0162 D10	2.2	0.18	0.082
437	55032	R0164 F1	2.72	0.256	0.094
438	55036	R0165 E2	3.51	0.279	0.079
439	55039	R0165 G5/LST-S5	3.14	0.195	0.062

The ratio of signal 1 to signal 2 in the table above provides a measure of the level of expression of the identified sequences in tumor versus normal tissues. For example, for SEQ ID NO: 422, the tumor-specific signal was 2.35 times that of the signal for the normal tissues tested; for SEQ ID NO: 423, the tumor-specific signal was 52.52 times that of the signal for normal tissues, etc.

Additional analyses were performed on lung microarray chips containing sequences from the LST-S5 and MS1 subtracted libraries. In one analysis, using a criteria of greater than or equal to 2-fold overexpression in tumors and an average expression in normal tissues less than or equal to 0.2, the following results were obtained and are described in Table 3:

Table 3

SEQ ID NO:	Ref No:	Element (96)	Ratio	Median Signal 1	Median Signal 2	Library
440	56710.1	R0121 E12	5.26	0.804	0.153	Mets3209-S1
441	56712.1	R0121 F7	2.82	0.453	0.161	Mets3209-S1
442	56716.1	R0159 G12	2.44	0.414	0.17	LST-S5
443	56718.1	R0160 A4	5.99	1.07	0.178	LST-S5
444	56723.1	R0163 A12	4.28	0.571	0.133	LST-S5
445	56724.1	R0164 C2	2.79	0.312	0.112	LST-S5
446	56730.1	R0164 G3	2.54	0.314	0.123	LST-S5
447	56732.1	R0165 G10	4.0	0.882	0.221	LST-S5

In another analysis, visual analysis was used for identifying cDNAs over-expressed in selected tumor samples. Some of these cDNAs were found to be preferentially over-expressed in small cell lung carcinoma samples, even though the original cDNAs were identified from subtracted non-small cell lung carcinoma tumor samples. The results of this analysis are summarized in Table 4 below.

Table 4

SEQ ID NO:	Ref No:	Element (96)	Ratio	Median Signal 1	Median Signal 2	Library
448	58375.3	R0164 H1	-	-	-	LST-S5
449	60982.1	R0160 G8	10.7	0.807	0.075	LST-S5
450	60983.2	R0160 E3	4.78	0.309	0.065	LST-S5

QUANTITATIVE REAL-TIME RT-PCR ANALYSIS OF LSCC AND ADENOCARCINOMA-SPECIFIC GENES

Quantitation of PCR product relies on the few cycles where the amount of DNA amplifies logarithmically from barely above the background to the plateau. Using continuous fluorescence monitoring, the threshold cycle number where DNA amplifies logarithmically is easily determined in each PCR reaction. There are two fluorescence detecting systems. One is based upon a double-strand DNA specific binding dye SYBR Green I dye. The other uses TaqMan probe containing a Reporter dye at the 5' end (FAM) and a Quencher dye at the 3' end (TAMRA) (Perkin Elmer/Applied Biosystems Division, Foster City, CA). Target-specific PCR amplification results in cleavage and release of the Reporter dye from the Quencher-containing probe by the nuclease activity of AmpliTaq Gold™ (Perkin Elmer/Applied Biosystems Division, Foster City, CA). Thus, fluorescence signal generated from released reporter dye is proportional to the amount of PCR product. Both detection methods have been found to generate comparable results To compare the relative level of gene expression in multiple tissue samples, a panel of cDNAs is constructed using RNA from tissues and/or cell lines, and real-time PCR is performed using gene specific primers to quantify the copy number in each cDNA sample. Each cDNA sample is generally performed in duplicate and each reaction repeated in duplicated plates. The final Real-time PCR result is typically reported as an average of copy number of a gene of interest normalized against internal actin number in each cDNA sample. Real-time PCR reactions may be performed on a GeneAmp 5700 Detector using SYBR Green I

dye or an ABI PRISM 7700 Detector using the TaqMan probe (Perkin Elmer/Applied Biosystems Division, Foster City, CA).

EXAMPLE 2

L587S FULL-LENGTH cDNA AND PROTEIN

Full-length cDNA for L587S was obtained. The cDNA encodes a novel protein with 255 amino acids. L587S demonstrated over-expression in lung small cell carcinoma by microarray, real-time PCR, and Northern analysis. The full-length cDNA is set forth in SEQ ID NO:453 and represents an extended sequence of clone 55022 (SEQ ID NO:435). The L587S amino acid sequence is set forth in SEQ ID NO:454. Microarray analysis, carried out essentially as described in example 1 above, demonstrated that L587S is overexpressed in small cell lung carcinoma tumors relative to normal tissues. By Real time PCR, L587 was found to be highly expressed in all of the small cell primary tumors and tumor cell lines that were tested. The expression levels in the small cell primary tumors and tumor cell lines were typically from about 5-fold to greater than 50-fold higher than those observed in normal lung tissues. Expression was also detected in adenocarcinoma and squamous lung tumor pools. No significant expression was observed in normal lung, brain, pituitary gland, adrenal gland, thyroid gland, pancreas, heart, liver, skeletal muscle, kidney, small intestine, bladder, skin, salivary gland, PBMC, spleen or spinal cord. Some low level expression was observed in stomach, colon, esophagus, trachea, bone marrow, lymph node and thymus, however this expression was at a level much less than was observed in the small cell tumors and tumor cell lines. Northern analysis of L587S demonstrated the presence of 2 isoforms of about 2 kb in lung small cell carcinoma.

EXAMPLE 3

EXPRESSION IN *E. COLI* OF A L587S HIS TAG FUSION PROTEIN

The full length cDNA sequence of L587S (SEQ ID NO:453) was described in Example 2. It was found to be highly overexpressed in tumor tissue compared to normal tissue. This example describes the expression L587S in E. coli.

PCR was performed on the L587S coding region with the following primers:

Forward primer PDM-647: 5' gcctcgtcagatctggaacaattatgctc 3' (SEQ ID NO:455) Tm 61°C.

Reverse primer PDM-648: 5' cgtaactcgagtcatcaggttataacataac 3' (SEQ ID NO:456) TM 59°C.

The PCR conditions were as follows:

10µl 10X Pfu buffer

1.0µl 10mM dNTPs

2.0µl 10µM each primer

83µl sterile water

1.5µl Pfu DNA polymerase (Stratagene, La Jolla, CA)

50ηg DNA

PCR amplification was carried out under the following conditions:

An initial 96°C for 2 minutes, followed by 40 cycles of 96°C for 20 seconds, 60°C for 15 seconds, and 72°C for 90 seconds. This was followed by a final 72°C extension step for 4 minutes.

The PCR product was digested with XhoI restriction enzyme, gel purified and cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and XhoI restriction enzymes. The correct construct was confirmed by DNA sequence analysis and then transformed into BLR (DE3) pLysS and BLR (DE3) CodonPlus RP cells for expression. Protein expression was induced using IPTG.

The amino acid sequence of expressed recombinant L587S is disclosed in SEQ ID NO:457, and the DNA coding region sequence is shown in SEQ ID NO:458.

EXAMPLE 4

SYNTHESIS OF POLYPEPTIDES

Polypeptides may be synthesized on a Perkin Elmer/Applied Biosystems Division 430A peptide synthesizer using FMOC chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

EXAMPLE 5

DETECTION OF L587S-SPECIFIC ANTIBODIES IN LUNG PLURAL EFFUSION (LPE) FROM PATIENTS WITH SMALL CELL LUNG CARCINOMAS (SCLC)

Recombinant protein was generated for L587S (SEQ ID NO: 457) and used in a protein based ELISA to detect the presence of L587S specific antibodies in the LPE of patients suffering from SCLC. Three of seven SCLC patients had detectable levels of L587S specific antibodies (patient #s: 298-42, 574-57, and G412), while Abs for L587S were undetectable in the 6 normal donors tested. This finding was confirmed by Western Blot analysis. L587S protein was run on an SDS-PAGE and probed with

the LPE from the seven patients suffering from SCLS. Consistent with data generated from the protein based ELISA, analysis showed the presence of a L587S specific band in the same patients that were positive using the protein based ELISA (patient #s: 298-42, 574-57, and G412).

To determine which portions of O587S were immunogenic, peptides specific for O587S were synthesized. These peptides were 15-mers that overlapped by 10 amino acids. Patients #574-57 and #298-42 were both tested using a peptide based ELISA. Epitope analysis revealed that patient #574-57 reacted against peptides #15 (amino acid 71-85) and #23 (amino acid (111-125), the sequences for which are disclosed in SEQ ID NOs:459 and 460). Patient #298-42 was shown to react against peptides #1 (amino acids 1-15), #9 (amino acids 41-55), and #45 (amino acids 221-235), the sequences for which are disclosed in SEQ ID NOs:461-463.

EXAMPLE 6

GENERATION OF L587S-SPECIFIC CYTOTOXIC T LYMPHOCYTES (CTL)

To determine if L587S is capable of generating a $CD8^+$ T cell immune response, CTLs were generated using *in vitro* priming methodologies. To do this, peripheral blood mononuclear cells (PBMC) were isolated from normal donors by Percol gradient followed by plastic adherence. The adherent population was then cultured for 5 days in the presence of RPMI medium supplemented with 1% human serum, 50ng/ml GM-CSF, and 30ng/ml of IL-4. After 5 days of culture the non-adherent cells, which constituted the dendritic cell (DC) population, were harvested and infected for 24 hours with L587S-expressing adenovirus at a multiplicity of infection (MOI) of 10. The DCs were then matured for an additional 24 hours by the addition of 2μg/ml of CD40 ligand. In order to generate a CTL line, autologous PBMC were isolated and $CD8^+$ T cells were enriched for by negative selection using magnetic beads conjugated to $CD4^+$, $CD14^+$, and $CD16^+$. $CD8^+$ T cell lines specific for L578S were established in round bottom 96-well plates using 10,000 L587S expressing DCs and 100,000 $CD8^+$ T cells per well in RPMI supplemented with 10% human serum, 5ng/ml IL-12, and 10ng/ml IL-6. The cultures were re-stimulated every 7 days using autologous fibroblasts that had been retrovirally transduced to express L587S and

CD80. The cells were also stimulated with IFN-gamma to upregulate MHC Class I. The media was supplemented with 10U/ml of IL-2 at the time of re-stimulation as well as on days 2 and 5 following stimulation. Following 4 cycles of stimulation, three L587S specific $CD8^+$ T cell lines were identified that produced IFN-gamma in response to exposure to IFN-gamma treated L587S/CD80 expressing autologous fibroblasts, but did not respond to cells transduced with a control antigen. These 3 lines were cloned in 96-well plates using a frequency of either 0.5 or 2 $CD8^+$ T cells/well in the presence of 75,000 irradiated PBMC, 10,000 irradiated B-LCL, 30ng/ml OKT3 (anti-CD3), and 50u/ml IL-2. After 2 weeks of cloning, an aliquot of cells were taken from wells positive for growth and these cells tested against L587S transduced fibroblasts. Elispot results showed that one clone, 5E9/A6, reacted specifically in response to fibroblasts expressing L587S.

EXAMPLE 7

IDENTIFICATION OF L587S IMMUNOGENIC PEPTIDES THAT ARE CAPABLE OF STIMULATING A CD4-SPECIFIC T HELPER CELL RESPONSE

A series of peptides derived from the L587S amino acid sequence were synthesized and used in *in vitro* priming experiments to generate $CD4^+$ T Helper cells specific for L587S. These peptides ranged in size from 19-22 mers that overlapped by 5 amino acids.

To generate the CD4+ T helper cells, peptides were combined into pools of 10, and pulsed onto DCs at a concentration of 0.25μg/ml for 24 hours. The DCs were then washed and mixed with positively selected $CD4^+$ T cells in round bottom 96-well plates. The cultures were re-stimulated weekly on fresh DC loaded with peptide pools. Following a total of 3 stimulations, the cells were rested for a week before being tested for specificity using antigen-presenting cells (APC) pulsed with each of the peptide pools. The specificity of the T cell lines was measured using an IFN-gamma ELISA and a T cell proliferation assay. To perform these assays, adherent monocytes loaded with either the relevant peptide pool or an irrelevant peptide pool were used as APC. T cell lines that specifically recognize an L587S-specific peptide pool, both by

cytokine release and proliferation were identified. T cells were found to react against peptide pools 1, 3, and 4.

CD4 T cell lines that tested positive for a specific peptide pool, were then screened against the individual peptides from that pool. For these assays, APC were pulsed with 0.25μg of pooled L587S peptides or 0.25μg of individual peptides. Peptides capable of generating a $CD4^+$ T helper responses in the donors tested are summarized in Table 5.

Table 5

Line /Peptide Pool Positive	Prolif. in response to pool (SI)	IFN-γ production in response to pool	Specific Peptide (aa)	Prolif. In response to specific peptide (SI)	IFN-γ in response to specific peptide	SEQ ID NO
1A3/1	52	41	16-35	46	30	472
1C11/1	7.6	9	36-55	6.8	7	471
1C11/1	7.6	9	41-60	4.8	6	470
1H8/1	212	44	11-30	148	21	473
1H8/1	212	44	16-35	116	16	472
1E4/1	2.2	3.3	36-55	2.3	3.6	471
1E4/1	2.2	3.3	41-60	32	3.8	470
3D6/3	47	7.3	146-165	40	6.6	469
4A3/4	4.3	9.6	161-180	2.9	8	466
4F3/4	132	38	151-570	99	27	468
4F3/4	132	38	156-175	50	4.4	465
4F3/4	132	38	166-185	63	14	467
4F3/4	132	38	171-190	88	36	464

Prolif=proliferation; aa=amino acids; SI=stimulation index

From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration,

various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

What is claimed:

1. An isolated polynucleotide comprising a sequence selected from the group consisting of:

a. sequences provided in SEQ ID NO: 1-451, 453, and 458;

b. complements of the sequences provided in SEQ ID NO: 1-451, 453, and 458;

c. sequences consisting of at least 20 contiguous residues of a sequence provided in SEQ ID NO: 1-451, 453, and 458;

d. sequences that hybridize to a sequence provided in SEQ ID NO: 1-451, 453, and 458, under moderately stringent conditions;

e. sequences having at least 75% identity to a sequence of SEQ ID NO: 1-451, 453, and 458;

f. sequences having at least 90% identity to a sequence of SEQ ID NO: 1-451, 453, and 458; and

g. degenerate variants of a sequence provided in SEQ ID NO: 1-451, 453, and 458.

2. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

a. sequences encoded by a polynucleotide of claim 1; and

b. sequences having at least 70% identity to a sequence encoded by a polynucleotide of claim 1; and

c. sequences having at least 90% identity to a sequence encoded by a polynucleotide of claim 1.

d. SEQ ID NOs:452, 454, 457, and 459-473;

e. sequences having at least 70% identity to a sequence encoded by SEQ ID NOs:452, 454, 457, and 459-473; and

f. sequences having at least 90% identity to a sequence encoded by SEQ ID NOs:452, 454, 457, and 459-473.

3. An expression vector comprising a polynucleotide of claim 1 operably linked to an expression control sequence.

4. A host cell transformed or transfected with an expression vector according to claim 3.

5. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a polypeptide of claim 2.

6. A method for detecting the presence of a cancer in a patient, comprising the steps of:

a. obtaining a biological sample from the patient;

b. contacting the biological sample with a binding agent that binds to a polypeptide of claim 2;

c. detecting in the sample an amount of polypeptide that binds to the binding agent; and

d. comparing the amount of polypeptide to a predetermined cut-off value and therefrom determining the presence of a cancer in the patient.

7. A fusion protein comprising at least one polypeptide according to claim 2.

8. An oligonucleotide that hybridizes to a sequence recited in SEQ ID NO: 1-451, 453, and 458 under moderately stringent conditions.

9. A method for stimulating and/or expanding T cells specific for a tumor protein, comprising contacting T cells with at least one component selected from the group consisting of:

a. polypeptides according to claim 2;

b. polynucleotides according to claim 1; and

c. antigen-presenting cells that express a polypeptide according to claim 2,

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

10. An isolated T cell population, comprising T cells prepared according to the method of claim 9.

11. A composition comprising a first component selected from the group consisting of physiologically acceptable carriers and immunostimulants, and a second component selected from the group consisting of:

a. polypeptides according to claim 2;

b. polynucleotides according to claim 1;

c. antibodies according to claim 5;

d. fusion proteins according to claim 7;

e. T cell populations according to claim 10; and

f. antigen presenting cells that express a polypeptide according to claim 2.

12. A method for stimulating an immune response in a patient, comprising administering to the patient a composition of claim 11.

13. A method for the treatment of a cancer in a patient, comprising administering to the patient a composition of claim 11.

14. A method for determining the presence of a cancer in a patient, comprising the steps of:

a. obtaining a biological sample from the patient;

b. contacting the biological sample with an oligonucleotide according to claim 8;

c. detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

d. compare the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence of the cancer in the patient.

15. A diagnostic kit comprising at least one oligonucleotide according to claim 8.

16. A diagnostic kit comprising at least one antibody according to claim 5 and a detection reagent, wherein the detection reagent comprises a reporter group.

17. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

a. incubating CD4+ and/or CD8+ T cells isolated from a patient with at least one component selected from the group consisting of: (i) polypeptides according to claim 2; (ii) polynucleotides according to claim 1; and (iii) antigen presenting cells that express a polypeptide of claim 2, such that T cell proliferate;

b. administering to the patient an effective amount of the proliferated T cells,

and thereby inhibiting the development of a cancer in the patient.

18. The fusion protein of claim 7, wherein the fusion protein comprises an amino acid sequence as provided in SEQ ID NO:457.

SEQUENCE LISTING

<110> Corixa Corporation
Wang, Tongtong
McNeill, Patricia D.
Watanabe, Yoshihiro
Carter, Darrick
Henderson, Robert A.
Kalos, Michael D.

<120> COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER

<130> 210121.539PC

<140> PCT
<141> 2001-06-28

<160> 473

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 147
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 4, 18, 21, 24, 29, 35, 40, 46, 49, 69, 92, 121, 133
<223> n = A,T,C or G

<400> 1
ttgngtattg ggcgccangg nggnttttnt tttcnccagn gagacnggnc aacagctgat 60
tgcccttcnc cgcctggccc tgagagagtt gnagcaagcg gtccacgctg gtttgcccca 120
ncaggcgaaa atnctgtttg atggtgg 147

<210> 2
<211> 595
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 565, 572
<223> n = A,T,C or G

<400> 2
ctagtattaa taggcttaat aattgttggc aaggatcctt ttgctttctt tggcatgcaa 60
gctcctagca tctggcagtg gggccaagaa aataaggttt atgcatgtat gatggttttc 120
ttcttgagca acatgattga gaaccagtgt atgtcaacag gtgcatttga gataacttta 180
aatgatgtac ctgtgtggtc taagctggaa tctggtcacc ttccatccat gcaacaactt 240
gttcaaattc ttgacaatga aatgaagctc aatgtgcata tggattcaat cccacaccat 300
cgatcatagc accacctatc agcactgaaa actcttttgc attaagggat cattgcaaga 360

gcagcgtgac tgacattatg aaggcctgta ctgaagacag caagctgtta gtacagacca 420
gatgctttct tggcaggctc gttgtacctc ttggaaaacc tcaatgcaag atagtgtttc 480
agtgctggca tattttggaa ttctgcacat tcatggagtg caataatact gtatagcttt 540
ccccacctcc cacaaaatca cccanttaat gngtgtgtgt gtgttttttt taagg 595

<210> 3
<211> 553
<212> DNA
<213> Homo sapiens

<400> 3
ctagtccagt gtggtggaat tcattttgac tgagcaaccc tagtgacagg agccgaagca 60
gcagcgcagg ttgtccccgt ttcccctccc ccttcccttc tccggttgcc ttcccgggcc 120
ccttacactc cacagtcccg gtcccgccat gtcccagaaa caagaagaag agaaccctgc 180
ggaggagacc ggcgaggaga agcaggacac gcaggagaaa gaaggtattc tgcctgagag 240
agctgaagag gcaaagctaa aggccaaata cccaagccta ggacaaaagc ctggaggctc 300
cgacttcctc atgaagagac tccagaaagg gcaaaagtac tttgactcag gagactacaa 360
catggccaaa gccaagatga agaataagca gctgccaagt gcaggaccag acaagaacct 420
ggtgactggt gatcacatcc ccaccccaca ggatctgccc cagagaaagt cctcgctcgt 480
caccagcaag cttgcgggtg gccaagttga atgatgctgc ccggggctct gccagatcct 540
gagacgcttc cct 553

<210> 4
<211> 494
<212> DNA
<213> Homo sapiens

<400> 4
ctagtccagt gtggtggaat tcattttgac tgagcaaccc tagtgacagg agccgaagca 60
gcagcgcagg ttgtccccgt ttcccctccc ccttcccttc tccggttgcc ttcccgggcc 120
ccttacactc cacagtcccg gtcccgccat gtcccagaaa caagaagaag agaaccctgc 180
ggaggagacc ggcgaggaga agcaggacac gcaggagaaa gaaggtattc tgcctgagag 240
agctgaagag gcaaagctaa aggccaaata cccaagccta ggacaaaagc ctggaggctc 300
cgacttcctc atgaagagac tccagaaagg gcaaaagtac tttgactcag gagactacaa 360
catggccaaa gccaagatga agaataagca gctgccaagt gcaggaccag acaagaacct 420
ggtgactggt gatcacatcc ccaccccaca ggatctgccc cagagaaagt cctcgctcgt 480
caccagcaag cttg 494

<210> 5
<211> 63
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 53
<223> n = A,T,C or G

<400> 5
ctagtccagt gtggtggaat tcccaggccc tggaccgcca aacagctact canctgctta 60
agc 63

<210> 6
<211> 357
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> 14
<223> n = A,T,C or G

<400> 6
ctagtccagt gtgntggaat tcgaccagca ccatggcggt tggcaagaac aagcgcctta 60
cgaaaggcgg caaaaaggga gccaagaaga aagtggttga tccattttct aagaaagatt 120
ggtatgatgt gaaagcacct gctatgttca atataagaaa tattggaaag acgctcgtca 180
ccaggaccca aggaaccaaa attgcatctg atggtctcaa gggtcgtgtg tttgaagtga 240
gtcttgctga tttgcagaat gatgaagttg catttagaaa attcaagctg attactgaag 300
atgttcaggg taaaaactgc ctgactaact tccatggcat ggatcttacc cgtgaca 357

<210> 7
<211> 442
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 14, 15
<223> n = A,T,C or G

<400> 7
ctagtccagt gtgnnggaat tccgcggcgg caagatggca gtgcaaatat ccaagaagag 60
gaagtttgtc gctgatggca tcttcaaagc tgaactgaat gagtttctta ctcgggagct 120
ggctgaagat ggctactctg gagttgaggt gcgagttaca ccaaccagga cagaaatcat 180
tatcttagcc accagaacac agaatgttct tggtgagaag ggccggcgga ttcgggaact 240
gactgctgta gttcagaaga ggtttggctt tccagagggc agtgtagagc tttatgctga 300
aaaggtggcc actagaggtc tgtgtgccat tgcccaggca gagtctctgc gttacaaact 360
cctaggaggg cttgctgtgc ggagggcctg ctatggtgtg ctgcggttca tcatggagag 420
tggggccaaa ggctgcgagg tt 442

<210> 8
<211> 108
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 7, 12, 75, 81, 95
<223> n = A,T,C or G

<400> 8
ctgcttnaac antgcttgga cggaacccgg cgctcgttcc ccaccccggc cggccgccca 60
tagccagccc tccgncacct nttcaccgca ccctnggact gccccaag 108

<210> 9
<211> 546
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 13
<223> n = A,T,C or G

<400> 9
ctagtccagt gtngtggaat tccttggttc cgcgttccct gcacaaaatg cccggcgaag 60

ccacagaaac cgtccctgct acagagcagg agttgccgca gccccaggct gagacagggt 120
ctggaacaga atctgacagt gatgaatcag taccagagct tgaagaacag gattccaccc 180
aggcaaccac acaacaagcc cagctggcgg cagcagctga aatcgatgaa gaaccagtca 240
gtaaagcaaa acagagtcgg agtgaaaaga aggcacggaa ggctatgtcc aaactgggtc 300
ttcggcaggt tacaggagtt actagagtca ctatccggaa atctaagaat atcctctttg 360
tcatcacaaa accagatgtc tacaagagcc ctgcttcaga tacttacata gtttttgggg 420
aagccaagat cgaagattta tcccagcaag cacaactagc agctgctgag aaattcaaag 480
ttcaaggtga agctgtctca aacattcaag aaaacacaca gactccaact gtacaagagg 540
agagtg 546

<210> 10
<211> 426
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 2, 11, 26, 197, 341
<223> n = A,T,C or G

<400> 10
gnaattcgtt ntttggttcc tgcgtnggga ttccgtgtac aatccataga catctgacct 60
cggcacttag catcatcaca gcaaactaac tgtagccttt ctctctttcc ctgtagaaac 120
ctctgcgcca tgagagccaa gtggaggaag aagcgaatgc gcaggctgaa gcgcaaaaga 180
agaaagatga ggcagangtc caagtaaacc gctagcttgt tgcaccgtgg aggccacagg 240
agcagaaaca tggaatgcca gacgctgggg atgctggtac aagttgtggg actgcatgct 300
actgtctaga gcttgtctca atggatctag aacttcatcg ncctctgatc gccgatcacc 360
tctgagaccc accttgctca taaacaaaat gcccatgttg gtcctctgcc ctggacctgt 420
gacatt 426

<210> 11
<211> 416
<212> DNA
<213> Homo sapiens

<400> 11
ctagtttaag gagactggcc gaagctctgc ccaaacaatc tgtggatgga aaagcaccac 60
ttgctactgg agaggatgat gatgatgaag ttccagatct tgtggagaat tttgatgagg 120
cttccaagaa tgaggcaaac tgaattgagt caacttctga agataaaacc tgaagaagtt 180
actgggagct gctattttat attatgactg ctttttaaga aatttttgtt tatggatctg 240
ataaaatcta gatctctaat atttttaagc ccaagcccct tggacactgc agctctttc 300
agtttttgct tatacacaat tcattctttg cagctaatta agccgaagaa gcctgggaat 360
caagtttgaa acaaagatta ataaagttct ttgcctagta aaaaaaaaaa aaaaaa 416

<210> 12
<211> 59
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 7, 22, 57
<223> n = A,T,C or G

<400> 12
ctagtcnagt gtggtggaat tncaaagaac tgggtactaa acactgagca gatctgntc 59

<210> 13

<211> 474
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 397, 435
<223> n = A,T,C or G

<400> 13
ctagtgacta attttccctt acagttcctg cttggtccca cccactgaag tagctcatcg 60
tagtgcgggc cgtattagaa gcagtggggt acgttagact cagatggaaa agtattctag 120
gtgccagtgt taggatgtca gttttacaaa ataatgaagc aattagctat gtgattgaga 180
gttattgttt ggggatgtgt gttgtggttt tgctttttt ttttagactg tattaataaa 240
catacaacac aagctggcct tgtgttgctg gttcctattc agtatttcct ggggattgtt 300
tgctttttaa gtaaaacact tctgacccat agctcagtat gtctgaattc cagaggtcac 360
atcagcatct ttctgctttg aaaactctca cagctgnggc tgcttcactt agatgcagtg 420
agacacatag ttggngttcc gattttcaca tccttccatg tatttatctt gaag 474

<210> 14
<211> 186
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 60, 171
<223> n = A,T,C or G

<400> 14
ttacagttcc tgcttggtcc cacccactga agtagctcat cgtagtgcgg gccgtattan 60
aagcagtggg gtacgttaga ctcagatgga aaagtattct aggtgccagt gttaggatgt 120
cagttttaca aaataatgaa gcaattagct atgtgattga gagttattgg nttggggatg 180
tgtgtt 186

<210> 15
<211> 456
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 441
<223> n = A,T,C or G

<400> 15
cgggagagag gccgagatgg cagatgagat tgccaaggct caggtcgctc ggcctggtgg 60
cgacacgatc tttgggaaga tcatccgcaa ggaaatacca gccaaaatca tttttgagga 120
tgaccggtgc cttgctttcc atgacatttc ccctcaagca ccaacacatt ttctggtgat 180
acccaagaaa catatatccc agatttctgt ggcagaagat gatgatgaaa gtcttcttgg 240
acacttaatg attgttggca agaaatgtgc tgctgatctg ggcctgaata agggttatcg 300
aatggtggtg aatgaaggtt cagatggtgg acagtctgtc tatcacgttc atctccatgt 360
tcttggaggt cggcaaatgc attggcctcc tggttaagca cgttttgggg ataattttct 420
cttctttagg caatgattaa nttaggcaat ttccag 456

<210> 16
<211> 495
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 470, 484, 485, 486
<223> n = A,T,C or G

<400> 16
ctagtccagt gtggnggaat tcgccgaaat gggcaagttc atgaaacctg ggaaggtggt 60
gcttgtcctg gctggacgct actccggacg caaagctgtc atcgtgaaga acattgatga 120
tggcacctca gatcgcccct acagccatgc tctggtggct ggaattgacc gctacccccg 180
caaagtgaca gctgccatgg gcaagaagaa gatcgccaag agatcaaaga taaaatcttt 240
tgtgaaagtg tataactaca atcacctaat gcccacaagg tactctgtgg atatcccctt 300
ggacaaaact gtcgtcaata aggatgtctt cagagatcct gctcttaaac gcaaggcccg 360
acgggaggcc aaggtcaagt ttgaagagag atacaagaca ggcaagaaca agtggttctt 420
ccagaaactg cggttttaga tgctttgttt tgatcattaa aaattataan gaaaaaaaaa 480
aaannnaaaa agggc 495

<210> 17
<211> 264
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 14
<223> n = A,T,C or G

<400> 17
ctagtccagt gtgntggaat tcattagaca ctttggaaga tggcataacc tgtctcacct 60
ggacttaagc gtctggctct aattcacagt gctcttttct cctcactgta tccaggttcc 120
ctcccagagg agccaccagt tctcatgggt ggcactcagt ctctcttctc tccagctgac 180
taaacttttt ttctgtacca gttaattttt ccaactacta atagaataaa ggcagttttc 240
taaaaaaaaa aaaaaaaaaa gggc 264

<210> 18
<211> 512
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 13, 120, 284, 313
<223> n = A,T,C or G

<400> 18
ctagtttcca aancggagac ttccgacttc cttacaggat gaggctgggc attgcctggg 60
acagcctatg taaggccatg tgccccttgc cctaacaact cactgcagtg ctcttcatan 120
acacatcttg cagcattttt cttaaggcta tgcttcagtt tttctttgta agccatcaca 180
agccatagtg gtaggtttgc cctttggtac agaaggtgag ttaaagctgg tggaaaaggc 240
ttattgcatt gcattcagag taacctgtgt gcatactcta gaanagtagg gaaaataatg 300
cttgttacaa ttngacctaa tatgtgcatt gtaaaataaa tgccatattt caaacaaaac 360
acgtaatttt tttacagtat gttttattac cttttgatat ctgttgttgc aatgttagtg 420
atgttttaaa atgtgatcga aaatataatg cttctaagaa ggaacagtag tggaatgaat 480
gtctaaaaga tctttatgtg tttatggtct gc 512

<210> 19
<211> 171

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 11, 18, 105, 158
<223> n = A,T,C or G

<400> 19
tcatactccc nggtgtantg tattctctaa aagctttaaa tgtctgcatg cagccagcca 60
tcaaatagtg aatggtctct ctttggctgg aattacaaaa ctcanagaaa tgtgtcatca 120
ggagaacatc ataacccatg aaggataaaa gccccaantg gtggtaactg a 171

<210> 20
<211> 205
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 42, 96, 100, 105, 140, 154, 156
<223> n = A,T,C or G

<400> 20
aattcatctg tgaaaatggt tcgctattca cttgacccgg anaaccccac gaaatcatgc 60
aaatcaagag gttccaatct tcgtgttcac tttaanaacn ctcgngaaac tgctcaggcc 120
atcaagggta tgcatatacn aaaagccacg aagnanctga aagatgtcac tttacagaaa 180
cagtgtgtac cattccgacg ttaca 205

<210> 21
<211> 600
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 583
<223> n = A,T,C or G

<400> 21
ctagtagtca tactccctgg tgtagtgtat tctctaaaag ctttaaatgt ctgcatgcag 60
ccagccatca aatagtgaat ggtctctctt tggctggaat tacaaaactc agagaaatgt 120
gtcatcagga gaacatcata acccatgaag gataaaagcc ccaaatggtg gtaactgata 180
atagcactaa tgctttaaga tttggtcaca ctctcaccta ggtgagcgca ttgagccagt 240
ggtgctaaat gctacatact ccaactgaaa tgttaaggaa gaagatagat ccaattaaaa 300
aaaattaaaa ccaatttaaa aaaaaaaaga acacaggaga ttccagtcta cttgagttag 360
cataatacag aagtcccctc tactttaact tttacaaaaa agtaacctga actaatctga 420
tgttaaccaa tgtatttatt tctgtggttc tgtttccttg ttccaatttg acaaaaccca 480
ctgttcttgt attgtattgc ccagggggag ctatcactgt acttgtagag tggtgctgct 540
ttaattcata aatcacaaat aaaagccaat tagctctata aanaaaaaaa aaaaaaaaaa 600

<210> 22
<211> 443
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> 165, 258, 280, 284, 299, 309, 331, 336, 343, 348, 369, 371,
380, 385, 393, 417, 422, 430
<223> n = A,T,C or G

<400> 22
ctagtccagt gtggtggaat tcgcagacca gacttcgctc gtactcgtgc gcctcgcttc 60
gcttttcctc cgcaaccatg tctgacaaac ccgatatggc tgagatcgag aaattcgata 120
agtcgaaact gaagaagaca gagacgcaag agaaaaatcc actgncttcc aaagaaacga 180
ttgaacagga gaagcaagca ggcgaatcgt aatgaggcgt gcgccgccaa tatgcactgt 240
acattccaca agcattgnct tcttatttta cttcttttan ctgnttaact ttgtaagang 300
caaagaggnt ggatcaagtt taaatgactg ngctgnccct ttnacatnaa agaactactg 360
acaacgaang ncgcgcctgn ctttnccatc tgnctatcta tctggctggc agggaangaa 420
anaacttgcn tgttggtgaa aga 443

<210> 23
<211> 506
<212> DNA
<213> Homo sapiens

<400> 23
ctagtccagt gtggtggaat tccgggtgtg ctctttgtga aattccacca tggcgtaccg 60
tggccagggt cagaaagtgc agaaggttat ggtgcagccc atcaacctca tcttcagata 120
cttacaaaat agatcgcgga ttcaggtgtg gctctatgag caagtgaata tgcggataga 180
aggctgtatc attggttttg atgagtatat gaaccttgta ttagatgatg cagaagagat 240
tcattctaaa acaaagtcaa gaaaacaact gggtcggatc atgctaaaag gagataatat 300
tactctgcta caaagtgtct ccaactagaa atgatcaatg aagtgagaaa ttgttgagaa 360
ggatacagtt tgtttttaga tgtcctttgt ccaatgtgaa catttattca tattgttttg 420
attaccctcg tgttactaca agatggcaat aaatactatg ggattgtttg tattaaaaaa 480
ttaaaaaaaa aaaaaaaaaa aagggc 506

<210> 24
<211> 490
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 445
<223> n = A,T,C or G

<400> 24
ctagtccagt gtggtggaat tcaagaactg ggtactcaac actgagcaga tctgttcttt 60
gagctaaaaa ccatgtgctg taccaagagt ttgctcctgg ctgctttgat gtcagtgctg 120
ctactccacc tctgcggcga atcagaagca agcaactttg actgctgtct tggatacaca 180
gaccgtattc ttcatcctaa atttattgtg ggcttcacac ggcagctggc caatgaaggc 240
tgtgacatca atgctatcat ctttcacaca aagaaaaagt tgtctgtgtg cgcaaatcca 300
aaacagactt gggtgaaata tattgtgcgt ctcctcagta aaaaagtcaa gaacatgtaa 360
aaactgtggc ttttctggaa tggaattgga catagcccaa gaacagaaag aaccttgctg 420
gggttggagg tttcacttgc acatnatgga gggtttagtg cttatctaat ttgtgcctca 480
cttggacttg 490

<210> 25
<211> 390
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> 1, 12, 13, 15, 34, 45, 52, 53, 94, 107, 116, 145, 154, 181,
203, 204, 223, 225, 243, 271, 280, 331, 340, 348
<223> n = A,T,C or G

<400> 25
ntagtccagt gnngnggaat tcaagaactg ggtnctcaac actgngcaga tnngttcttt 60
gagctaaaaa ccatgtgctg taccaagagt ttgntcctgg ctgcttngat gtcagngctg 120
ctactccacc tctgcggcga atcanaagca agcnactttg actgctgtct tggatacaca 180
naccgtattc ttcatcctaa atnnattgtg ggcttcacac ggnanctggc caatgaaggc 240
tgngacatca atgctatcat ctttcacaca nagaaaaagn tgtctgtgtg cgcaaatcca 300
aaacagactt gggtgaaata tattgtgcgt ntcctcagtn aaaaagtnaa gaacatgtaa 360
aaactgtggc ttttctggaa tggaattgga 390

<210> 26
<211> 516
<212> DNA
<213> Homo sapiens

<400> 26
ctagtccagt gtggtggaat tccttttgtc tttccgtgga gctgtcgcca tgaaggtcga 60
gctgtgcagt tttagcgggt acaagatcta ccccggacac gggaggcgct acgccaggac 120
cgacgggaag gttttccagt ttcttaatgc gaaatgcgag tcggctttcc tttccaagag 180
gaatcctcgg cagataaact ggactgtcct ctacagaagg aagcacaaaa agggacagtc 240
ggaagaaatt caaaagaaaa gaacccgccg agcagtcaaa ttccagaggg ccattactgg 300
tgcatctctt gctgatataa tggccaagag gaatcagaaa cctgaagtta gaaaggctca 360
acgagaacaa gctatcaggg ctgctaagga agcaaaaaag gctaagcaag catctaaaaa 420
gactgcaatg gctgctgcta aggcacctac aaaggcagca cctaagcaaa agattgtgaa 480
gcctgtgaaa gtttcagctc cccgagttgg tggaaa 516

<210> 27
<211> 268
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 13, 58, 60, 134, 140, 212, 222, 223, 227, 242, 255, 265
<223> n = A,T,C or G

<400> 27
ctagtccagt gtngtggaat tcggttggca agaacaagcg ccttacgaaa ggcggcanan 60
agggagccaa gaagaaagtg gttgatccat tttctaagaa agattggtat gatgtgaaag 120
cacctgctat gttnaatatn agaaatattg gaaagacgct cgtcaccagg acccaaggaa 180
ccaaaattgc atctgatggt ctcaagggtc gngtgtttga anngagnctt gctgatttgc 240
anaatgatga agttncattt ataanatt 268

<210> 28
<211> 451
<212> DNA
<213> Homo sapiens

<400> 28
ctagtccagt gtggtggaat tcggcagccc tgtttacagt cacctggctg gtggggtggc 60
aggtgctctc tctgaattaa ccctttgaga gctggccagg actctggact gattacccca 120
gcctggggtg gcatccaggg gctctaggag gtacctttg ctcctcaccc tggatctctt 180
ttccttccac ccaggtttct gcaggtaatg gtggcagcag cctctcttac acaaacccag 240
cagtggcagc cacttctgcc aacttgtagg ggcacgtcgc ccgctgagct gagtggccag 300

ccagtgccat tccactccac tcaggttctt cagggccaga gcccctgcac cctgtttggg 360
ctggtgagct gggagttcag gtgggctgct cacagcctcc ttcagaggcc ccaccaattt 420
ctcggacact tctcagtgtg tggaagctca t 451

<210> 29
<211> 405
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 20, 21, 23, 252, 368, 377, 378
<223> n = A,T,C or G

<400> 29
ctagtccagt gtggtggaan ncnccatttt tttggaaacc tctgcgccat gagagccaag 60
tggaggaaga agcgaatgcg caggctgaag cgcaaaagaa gaaagatgag gcagaggtcc 120
aagtaaaccg ctagcttgtt gcaccgtgga ggccacagga gcagaaacat ggaatgccag 180
acgctgggga tgctggtaca agttgtggga ctgcatgcta ctgtctagag cttgtctcaa 240
tggatctaga anttcatcgc cctctgatcg ccgatcacct ctgagaccca ccttgctcat 300
aaacaaaatg cccatgttgg tcctctgccc tggacctgtg acattctgga ctatttctgt 360
gtttattngt ggccganngt aacaaccata taataaatca cctct 405

<210> 30
<211> 398
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 23, 33, 60, 63, 89, 90, 93, 104, 132, 135, 136, 146, 157,
170, 222, 250, 276, 313, 327, 381, 385, 392, 393
<223> n = A,T,C or G

<400> 30
ctagtccagt gtggtggaat tcnctcggag gangccaagg tgcaacttcc ttcggtcgtn 60
ccnaatccgg gttcatccga caccagccnn ctncaccatg ccgncgaagt tcgaccccaa 120
cgagatcaaa gncgnntacc tgaggngcac cggaggngaa gtcggtgccn cttctgccct 180
ggcccccaag atcggccccc tgggtctgtc tccaaaaaaa gntggtgatg acattgccaa 240
ggcaacgggn gactggaagg gcctgaggat tacagngaaa ctgaccattc agaacagaca 300
ggcccagatt gangtggtgc cttctgnctc tgccctgatc atcaaagccc tcaaggaacc 360
accaagagac agaaagaaac ngaanaacat tnnacaca 398

<210> 31
<211> 317
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 16, 23, 52, 307
<223> n = A,T,C or G

<400> 31
nattcttgct ccttgnggcc ctntcctaca ctctggccag agataccaca gncaaacctg 60
gagccaaaaa ggacacaaag gactctcgac ccaaactgcc ccagaccctc tccagaggtt 120
ggggtgacca actcatctgg actcagacat atgaagaagc tctatataaa tccaagacaa 180
gcaacaaacc cttgatgatt attcatcact tggatgagtg cccacacagt caagctttaa 240

agaaagtgtt tgctgaaaat aaagaaatcc agaaattggc agagcagttt gtcctcctca 300
atctggntta tgaaaca 317

<210> 32
<211> 115
<212> DNA
<213> Homo sapiens

<400> 32
tgtcgctgat ggcatcttca aagctgaact gaatgagttt cttactcggg agctggctga 60
agatggctac tctggagttg aggtgcgagt tacaccaacc aggacagaaa tcatt 115

<210> 33
<211> 520
<212> DNA
<213> Homo sapiens

<400> 33
ctagtggatt tgggaaaggt tcttaagtag atcctgagac tatttgcatg cttctgtcta 60
aatgataatt aaaaggaaat ttcatggatt aaaccatggg tttaatgcag caaggaaact 120
tacaatgtcc ctttatatat aacatgcatc ttgttttgga tttgtgtcat tttttaatat 180
agctgattga cttcacagaa agcagctttt ttgaattcta atacataggt gtatatttgg 240
tattagttat tttgagttct tttcaactta taacactgta tacagttatt tctaaagcac 300
agatgaaata agttctgcat atttttaaat aatcacagtt ccctgttata cagataatgt 360
tctcactacc cataatatgt aggaacattg tttctcctta gccgtagtat gcatacacct 420
atccatgttc attctgacat cctttgttgt ctttataatt catgtggtag ttacctataa 480
ataaaaacaa atatgcgtta aaaaaaaaaa aaaaaagggc 520

<210> 34
<211> 377
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 19, 20, 365
<223> n = A,T,C or G

<400> 34
ctagtccagt gtggcggann tccttgacga ggctgcggtg tctgctgcta ttctccgagc 60
ttcgcaatgc cgcctaagga cgacaagaag aagaaggacg ctggaaagtc ggccaagaaa 120
gacaaagacc cagtgaacaa atccgggggc aaggccaaaa agaagaagtg gtccaaaggc 180
aaagttcggg acaagctcaa taacttagtc ttgtttgaca aagctaccta tgataaactc 240
tgtaaggaag ttcccaacta taaacttata accccagctg tggtctctga gagactgaag 300
attcgaggct ccctggccag ggcagccctt caggagctcc ttagtaaagg acttatcaaa 360
ctggnttcaa agcacag 377

<210> 35
<211> 85
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 40, 41, 55, 63, 69, 70
<223> n = A,T,C or G

<400> 35

cggcaatgag ggccgcgtgt ctgtggaaaa catcaagcan nctgttgcaa tctgnccaca 60
aanaatccnn ctttgacatt atttt 85

<210> 36
<211> 564
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 479, 518, 542
<223> n = A,T,C or G

<400> 36
ctagtccagt gtggtggaat tcacagaagc cacctttttt cattctttca ttttaaaaaa 60
aagtgagata tccacattcc ataaaattca ccctttgaaa gtacacaatg caagttttta 120
atatattcac aagtttgttt aatccttacc actgtctaat tcaagagtat tatcattacc 180
ccaaaaagaa acccattagc agtcactccg cattctcacc ttcccccatt tcctcccaac 240
cactaagtga ttttctgtct ctatggattt gcatattctg gacattttat agaaatggaa 300
tcatgcaata tatgatcttt tgtgtctggt gtctttcaat gaacaatatt gtcagtcttc 360
atccacactg aagcttgtat cagtagtgag tgcttccttt ttatggcggc atactaatcc 420
attggatggc tatccgacat ttgttttatc tatgcatcaa ttgcagtgag cctggaggng 480
gaagactctg gttttttttag tgagcccttc aagaaggnac acatcctggt gagaggatga 540
anacaccgga gttcactgaa aggg 564

<210> 37
<211> 442
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 433
<223> n = A,T,C or G

<400> 37
ctagtagtca tactccctgg tgtagtgtat tctctaaaag ctttaaatgt ctgcatgcag 60
ccagccatca aatagtgaat ggtctctctt tggctggaat tacaaaactc agagaaatgt 120
gtcatcagga gaacatcata acccatgaag gataaaagcc ccaaatggtg gtaactgata 180
atagcactaa tgctttaaga tttggtcaca ctctcaccta ggtgagcgca ttgagccagt 240
ggtgctaaat gctacatact ccaactgaaa tgttaaggaa gaagatagat ccaattaaaa 300
aaaattaaaa ccaatttaaa aaaaaaaaga acacaggaga ttccagtcta cttgagttag 360
cataatacag aagtcccctc tactttaact tttacaaaaa agtaacctga actaatctga 420
tgttaaccaa tgnatttatt tc 442

<210> 38
<211> 434
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 20, 62, 299, 381, 384, 403, 416
<223> n = A,T,C or G

<400> 38
ctagtccagt gtggncggan ttcgtggtcg tagcggtggc ggaggaggcg ggtacgaatc 60
anctgcgggc ggagacatgg ccaacatcgc ggtgcagcga atcaagcggg agttcaagga 120

ggtgctgaag agcgaggaga cgagcaaaaa tcaaattaaa gtagatcttg tagatgagaa 180
ttttacagaa ttaagaggag aaatagcagg acctccagac acaccatatg aaggaggaag 240
ataccaacta gagataaaaa taccagaaac atacccattt aatcccccta aggtccggnt 300
tatcactaaa atatggcatc ctaatattag ttccgtcaca ggggctattt gtttggatat 360
cctgaaagat caatgggcag ntgnaatgac tctccgcacg gtnttattgt cattgnaagc 420
actattggca gctg 434

<210> 39
<211> 573
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 23, 444, 495, 506, 509, 510, 554
<223> n = A,T,C or G

<400> 39
ctagtccagt gtggtggaat tcnccgcgcc agtcgcctag caggtcctct accggcttat 60
tcctgtgccg gatcttcatc ggcacagggg ccactgagac gtttctgcct ccctctttct 120
tcctccgctc tttctcttcc ctctcgttta gtttgcctgg gagcttgaaa ggagaaagca 180
cggggtcgcc ccaaacccct tctgcttctg cccatcacaa gtgccactac cgccatgggc 240
ctcactatct cctccctctt ctcccgacta tttggcaaga agcagatgcg cattttgatg 300
gttggattgg atgctgctgg caagacaacc attctgtata aactgaagtt aggggagata 360
gtcaccacca ttcctaccat tggttttaat gtggaaacag tagaatataa gaacatttgt 420
ttcacagtat gggatgttgg tggncaagat agaattaggc ctctctggaa gcattacttc 480
cagaataccc agggncttat ttttgnggnn aggatagcaa cgatcgtgaa agaattcagg 540
aagtagcaga tganctgcag aaaatgcttc tgg 573

<210> 40
<211> 247
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 8, 9, 11, 49, 131, 170, 235
<223> n = A,T,C or G

<400> 40
ggtggaannc nccacatatt ctatgattcc atttctatga agtgtgcana gtaggcaaat 60
ctataaagac atagattggt ggttgggggt tggggagtat aggaaatgac tcctgatggg 120
tacagggttt ntttgtggag tgatgaaagt gttctaaaat tgatggcggn aatggttgca 180
caactccata tgaaaaccac tgaattatat acactgtaaa tgggtgaatt gtatnggatg 240
tgaatta 247

<210> 41
<211> 523
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 500
<223> n = A,T,C or G

<400> 41
ctagtccagt gtggtggaat tcctttgagc taaaaaccat gtgctgtacc aagagtttgc 60

tcctggctgc tttgatgtca gtgctgctac tccacctctg cggcgaatca gaagcaagca 120
actttgactg ctgtcttgga tacacagacc gtattcttca tcctaaattt attgtgggct 180
tcacacggca gctggccaat gaaggctgtg acatcaatgc tatcatcttt cacacaaaga 240
aaaagttgtc tgtgtgcgca aatccaaaac agacttgggt gaaatatatt gtgcgtctcc 300
tcagtaaaaa agtcaagaac atgtaaaaac tgtggctttt ctggaatgga attggacata 360
gcccaagaac agaaagaacc ttgctggggt tggaggtttc acttgcacat catggagggt 420
ttagtgctta tctaatttgt gcctcactgg acttgtccaa ttaatgaagt tgattcatat 480
tgcatcatag tttgctttgn ttaagcatca cattaaagtt aaa 523

<210> 42
<211> 579
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 513, 517, 543
<223> n = A,T,C or G

<400> 42
ctagtccagt gtggtggaat tcctcgtctc aggccagttg cagccttctc agccaaacgc 60
cgaccaagga aaactcacta ccatgagaat tgcagtgatt tgcttttgcc tcctaggcat 120
cacctgtgcc ataccagtta aacaggctga ttctggaagt tctgaggaaa agcagcttta 180
caacaaatac ccagatgctg tggccacatg gctaaaccct gacccatctc agaagcagaa 240
tctcctagcc ccacagaatg ctgtgtcctc tgaagaaacc aatgacttta aacaagagac 300
ccttccaagt aagtccaacg aaagccatga ccacatggat gatatggatg atgaagatga 360
tgatgaccat gtggacagcc aggactccat tgactcgaac gactctgatg atgtagatga 420
cactgatgat tctcaccagt ctgatgagtc tcaccattct gatgaatctg atgaactggt 480
cactgatttt tccacggacc tgccagcaac cgnaagnttt cactccagtt gtccccacag 540
tangacacat atgatggccg aggtgatagt gtggtttat 579

<210> 43
<211> 404
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 388
<223> n = A,T,C or G

<400> 43
ctagtccagt gtggtggaat tccctattgt agatattgca ccctatgaca ttggtggtcc 60
tgatcaagaa tttggtgtgg acgttggccc tgtttgcttt ttataaacca aactctatct 120
gaaatcccaa caaaaaaaat ttaactccat atgtgttcct cttgttctaa tcttgtcaac 180
cagtgcaagt gaccgacaaa attccagtta tttatttcca aaatgtttgg aaacagtata 240
atttgacaaa gaaaaatgat acttctcttt ttttgctgtt ccaccaaata caattcaaat 300
gctttttgtt ttatttttt accaattcca atttcaaaat gtctcaatgg tgctataata 360
aataaacttc aacactcttt atgataanaa aaaaaaaaaa gggc 404

<210> 44
<211> 85
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 7, 27, 50

<223> n = A,T,C or G

<400> 44
cacatcnccg accaggtgag gtcccanctt gaagagaaag aaaacaagan gttccctgtg 60
tttaaggccg tgtcattcaa gaacc 85

<210> 45
<211> 428
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 19, 23, 24, 355, 424
<223> n = A,T,C or G

<400> 45
ctagtggtag cagtggaanc tcnnctaaaa atatctgggt tagtggactt tcatctaata 60
ccaaagctgc tgatttgaag aacctctttg gcaaatatgg aaaggttctg agtgcaaaag 120
tagttacaaa tgctcgaagt cctggggcaa aatgctatgg cattgtaact atgtcttcaa 180
gcacagaggt gtccaggtgt attgcacatc ttcatcgcac tgagctgcat ggacagctga 240
tttctgttga aaaagtaaaa ggtgatccct ctaagaaaga aatgaagaaa gaaaatgatg 300
aaaagagtag ttcaagaagt tctgggagat aaaaaaaata cgagtgatag aagtngcaag 360
acacaagcct ctgtcaaaaa agaagagaaa agatcgtctg agaaatctga aaaaaaaaaa 420
aaangggc 428

<210> 46
<211> 400
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 20, 23, 339, 352, 399
<223> n = A,T,C or G

<400> 46
ctagttgagg agtagaagan gangaccagc tagactccca tggaattgga actcctattc 60
cttgcttaga cattacaggt tatgctttga gatctctttg gggtgaagga ttgaaattaa 120
accctgagcc accgtgtcct tgtagagcac agagtagaga acaactggca gctttgaaaa 180
aacaccatga agaagaaatc gttcatcata agaaggagat tgagcgtctg cagaaagaaa 240
ttgagcgcca taagcagaag atcaaaatgc taaaacatga tgattaagtg cacaccgtgt 300
gccatagaat ggcacatgtc attgcccact tctgtgtana catggttctg gnttaactaa 360
tatttgtctg tgtgctacta acagattata ataaattgnc 400

<210> 47
<211> 437
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 19, 20, 112, 370
<223> n = A,T,C or G

<400> 47
ctagtagtca tactccctnn tgtagtgtat tctctaaaag ctttaaatgt ctgcatgcag 60
ccagccatca aatagtgaat ggtctctctt tggctggaat tacaaaactc anagaaatgt 120

gtcatcagga gaacatcata acccatgaag gataaaagcc ccaaatggtg gtaactgata 180
atagcactaa tgctttaaga tttggtcaca ctctcaccta ggtgagcgca ttgagccagt 240
ggtgctaaat gctacatact ccaactgaaa tgttaaggaa gaagatagat ccaattaaaa 300
aaaattaaaa ccaatttaaa aaaaaaaaga acacaggaga ttccagtcta cttgagttag 360
cataatacan gaagtcccct ctactttaac ttttacaaaa aaagtaacct gaactaatct 420
gatgttaacc aatgtat 437

<210> 48
<211> 451
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 440
<223> n = A,T,C or G

<400> 48
ctagtccagt gtggtggaat tctagatcgc catcatgaac gacaccgtaa ctatccgcac 60
tagaaagttc atgaccaacc gactacttca gaggaaacaa atggtcattg atgtccttca 120
ccccgggaag gcgacagtgc ctaagacaga aattcgggaa aaactagcca aaatgtacaa 180
gaccacaccg gatgtcatct ttgtatttgg attcagaact cattttggtg gtggcaagac 240
aactggcttt ggcatgattt atgattccct ggattatgca aagaaaaatg aacccaaaca 300
tagacttgca agacatggcc tgtatgagaa gaaaaagacc tcaagaaagc aacgaaagga 360
acgcaagaac agaatgaaga aagtcagggg gactgcaaag gccaatgttg gtgctggcaa 420
aaagccgaag gagtaaaggn gctgcaatga t 451

<210> 49
<211> 86
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 22, 28
<223> n = A,T,C or G

<400> 49
cggggtaggg gttggcgctc angcggcnac catggcgtat cacggcctca ctgtgcctct 60
cattgtgatg agcgtgttct ggggct 86

<210> 50
<211> 332
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 20, 23, 250, 281
<223> n = A,T,C or G

<400> 50
ctagtccagt gtggtggaan tcngcgagat ggcagtgcaa atatccaaga agaggaagtt 60
tgtcgctgat ggcatcttca aagctgaact gaatgagttt cttactcggg agctggctga 120
agatggctac tctggagttg aggtgcgagt tacaccaacc aggacagaaa tcattatctt 180
agccaccaga acacagaatg ttcttggtga gaagggccgg cggattcggg aactgactgc 240
tgtagttcan aagaggtttg gctttccaga gggcagtgta nagctttatg ctgaaaaggt 300
ggccactaga ggtctgtgtg ccattgccca gg 332

<210> 51
<211> 561
<212> DNA
<213> Homo sapiens

<400> 51
ctagtccagt gtggtggaat tcgaaggccc tgaagctgat ggggtcaaat gaaggtgaat 60
tcaaggctga aggaaatagc aaattcacct acacagttct ggaggatggt tgcacgaaac 120
acactgggga atggagcaaa acagtctttg aatatcgaac acgcaaggct gtgagactac 180
ctattgtaga tattgcaccc tatgacattg gtggtcctga tcaagaattt ggtgtggacg 240
ttggccctgt ttgcttttta taaaccaaac tctatctgaa atcccaacaa aaaaaattta 300
actccatatg tgttcctctt gttctaatct tgtcaaccag tgcaagtgac cgacaaaatt 360
ccagttattt atttccaaaa tgtttggaaa cagtataatt tgacaaagaa aaatgatact 420
tctcttttt tgctgttcca ccaaatacaa ttcaaatgct ttttgtttta ttttttacc 480
aattccaatt tcaaaatgtc tcaatggtgc tataataaat aaacttcaac actctttatg 540
ataaaaaaaa aaaaaaaggg c 561

<210> 52
<211> 295
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 19, 37, 66, 85, 183, 213, 226, 250
<223> n = A,T,C or G

<400> 52
gccgactcac acaaggcang tgggtgagga aatccanagt tgccatggag aaaattccag 60
tgtcancatt cttgctcctt gtggncctct cctacactct ggccagagat accacagtca 120
aacctggagc caaaaaggac acaaaggact ctcgacccaa actgccccag accctctcca 180
gangttgggg tgaccaactc atctggactc aanacatatg aagaanctct atataaatcc 240
aagacaagcn aacaaaccct tgatgattat tcatcacttg gatgagtgcc cacac 295

<210> 53
<211> 553
<212> DNA
<213> Homo sapiens

<400> 53
ctagtccagt gtggtggaat tcccaaagaa ctgggtactc aacactgagc agatctgttc 60
tttgagctaa aaaccatgtg ctgtaccaag agtttgctcc tggctgcttt gatgtcagtg 120
ctgctactcc acctctgcgg cgaatcagaa gcagcaagca actttgactg ctgtcttgga 180
tacacagacc gtattcttca tcctaaattt attgtgggct tcacacggca gctggccaat 240
gaaggctgtg acatcaatgc tatcatcttt cacacaaaga aaaagttgtc tgtgtgcgca 300
aatccaaaac agacttgggt gaaatatatt gtgcgtctcc tcagtaaaaa agtcaagaac 360
atgtaaaaac tgtggctttt ctggaatgga attggacata gcccaagaac agaaagaacc 420
ttgctggggt tggaggtttc acttgcacat catggagggt ttagtgctta tctaatttgt 480
gcctcactgg acttgtccaa ttaatgaagt tgattcatat tgcatcatag tttgctttgt 540
ttaagcatca cat 553

<210> 54
<211> 506
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> 487, 490
<223> n = A,T,C or G

<400> 54
ctagtccagt gtggtggaat tcgcatcttc tgaggtcaat taaaaggaga aaaaatacaa 60
tttctcactt tgcatttagt caaaagaaaa aatgctttat agcaaaatga aagagaacat 120
gaaatgcttc tttctcagtt tattggttga atgtgtatct atttgagtct ggaaataact 180
aatgtgtttg ataattagtt tagtttgtgg cttcatggaa actccctgta aactaaaagc 240
ttcagggtta tgtctatgtt cattctatag aagaaatgca aactatcact gtattttaat 300
atttgttatt ctctcatgaa tagaaattta tgtagaagca aacaaaatac ttttacccac 360
ttaaaaagag aatataacat tttatgtcac tataatcttt tgttttttaa gttagtgtat 420
attttgttgt gattatcttt ttgtggtgtg aataaatctt ttatcttgaa tgtaataaga 480
atttggnggn gtcaattgct tatttg 506

<210> 55
<211> 444
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 281, 402
<223> n = A,T,C or G

<400> 55
ctagtgacta attttccctt acagttcctg cttggtccca cccactgaag tagctcatcg 60
tagtgcgggc cgtattagaa gcagtggggt acgttagact cagatggaaa agtattctag 120
gtgccagtgt taggatgtca gttttacaaa ataatgaagc aattagctat gtgattgaga 180
gttattgttt ggggatgtgt gttgtggttt tgctttttt tttagactgt attaataaac 240
atacaacaca agctggcctt gtgttgctgg ttcctattca ntatttcctg gggattgttt 300
gctttttaag taaaacactt ctgacccata gctcagtatg tctgaattcc agaggtcaca 360
tcagcatctt tctgctttga aaactctcac agctgtggct gnttcactta gatgcagtga 420
gacacatagt tggtgttccg attt 444

<210> 56
<211> 247
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 65, 75, 88, 101, 103, 120, 196, 200, 237, 243
<223> n = A,T,C or G

<400> 56
ctgctattct ccgagcttcg caatgccgcc taaggacgac aagaagaaga aggacgctgg 60
aaagncggcc aaganagaca aagacccngt gaacaaatcc ngnggcaagg ccaaaaagan 120
gaagtggtcc aaaggcaaag ttcgggacaa gctcaataac ttagtcttgt ttgacaaagc 180
tacctatgat aaactntgtn aggaagttcc caactataaa cttataaccc cagctgnggt 240
ctntgag 247

<210> 57
<211> 475
<212> DNA
<213> Homo sapiens

<400> 57

ctagtccagt gtggtggaat tcatgtgccc aaccttcatg tcatgaaggc catgcagtct 60
ctcaagtccc gaggctacgt gaaggaacag tttgcctgga gacatttcta ctggtacctt 120
accaatgagg gtatccagta tctccgtgat taccttcatc tgccccccgga gattgtgcct 180
gccaccctac gccgtagccg tccagagact ggcaggcctc ggcctaaagg tctggagggt 240
gagcgacctg cgagactcac aagaggggaa gctgacagag atacctacag acggagtgct 300
gtgccacctg gtgccgacaa gaaagccgag gctggggctg ggtcagcaac cgaattccag 360
tttagaggcg gatttggtcg tggacgtggt cagccacctc agtaaaattg gagaggattc 420
ttttgcattg aataaactta cagccaaaaa accttaaaaa aaaaaaaaaa agggc 475

<210> 58
<211> 502
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 16, 19, 20
<223> n = A,T,C or G

<400> 58
ctagtccagt gtggtngann tccttttgtc tttccgtgga gctgtcgcca tgaaggtcga 60
gctgtgcagt tttagcgggt acaagatcta ccccggacac gggaggcgct acgccaggac 120
cgacgggaag gttttccagt ttcttaatgc gaaatgcgag tcggctttcc tttccaagag 180
gaatcctcgg cagataaact ggactgtcct ctacagaagg aagcacaaaa agggacagtc 240
ggaagaaatt caaaagaaaa gaacccgccg agcagtcaaa ttccagaggg ccattactgg 300
tgcatctctt gctgatataa tggccaagag gaatcagaaa cctgaagtta gaaaggctca 360
acgagaacaa gctatcaggg ctgctaagga agcaaaaaag gctaagcaag catctaaaaa 420
gactgcaatg gctgctgcta aggcacctac aaaggcagca cctaagcaaa agattgtgaa 480
gcctgtgaaa gtttcagctc cc 502

<210> 59
<211> 376
<212> DNA
<213> Homo sapiens

<400> 59
ctagttctgt gtgcctatga agttaatgct gcttattgtc tcattctgac ttcatggaga 60
attaatccca cctttaagca aaggctacta agttaatggt attttctgtg cagaaattaa 120
attttatttt cagcatttag cccaggaatt cttccagtag gtgctcagct atttaaaaac 180
aaaactattc tcaaacattc atcattagac aactggagtt tttgctggtt ttgtaaccta 240
ccaaaatgga taggctgttg aacattccac attcaaaagt tttgtagggt ggtgggaaat 300
gggggatctt caatgtttat tttaaaataa aataaaataa gttcttgact tttaaaaaaa 360
aaaaaaaaaa aagggc 376

<210> 60
<211> 356
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 346, 348, 351
<223> n = A,T,C or G

<400> 60
cttctacccg ggagctgtga cagtggcctg gaaggcagat ggcagccccg tcaaggcggg 60
agtggagacc accaaaccct ccaaacagag caacaacaag tacgcggcca gcagctacct 120
gagcctgacg cccgagcagt ggaagtccca cagaagctac agctgccagg tcacgcatga 180

agggagcacc gtggagaaga cagtggcccc tacagaatgt tcataggttc ccaactctaa 240
ccccacccac gggagcctgg agctgcagga tcccagggga ggggtctctc tccccatccc 300
aagtcatcca gcccttctcc ctgcactcat gaaaccccaa taaatntnct nattga 356

<210> 61
<211> 595
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 2, 18
<223> n = A,T,C or G

<400> 61
gntaagcttg atatcgantt cctgcagccc gggggatcca ctagtagtca gttgggagtg 60
gttgctatac cttgacttca tttatatgaa tttccacttt attaaataat agaaaagaaa 120
atcccggtgc ttgcagtaga gtgataggac attctatgct tacagaaaat atagccatga 180
ttgaaatcaa atagtaaagg ctgttctggc tttttatctt cttagctcat cttaaataag 240
cagtacactt ggatgcagtg cgtctgaagt gctaatcagt tgtaacaata gcacaaatcg 300
aacttaggat ttgtttcttc tcttctgtgt ttcgattttt gatcaattct ttaattttgg 360
aagcctataa tacagttttc tattcttgga gataaaaatt aaatggatca ctgatatttt 420
agtcattctg cttctcatct aaatatttcc atattctgta ttaggagaaa attaccctcc 480
cagcaccagc ccccctctca aacccccaac ccaaaaccaa gcattttgga atgagtctcc 540
tttagtttca gagtgtggat tgtataaccc atatactctt cgatgtactt gtttg 595

<210> 62
<211> 50
<212> DNA
<213> Homo sapiens

<400> 62
atcaattacg gggtcattag ttcatagccc atatatggag ttcctcgagt 50

<210> 63
<211> 422
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 404
<223> n = A,T,C or G

<400> 63
tacttcaatc attttcacag gcagccaaca agcaattaag agcagttata atagaggaag 60
ctgggggacc cattttgcac catgagtttg tgaaaaatct ggattaaaaa attacctctt 120
cagtgttttc tcatgcaaaa ttttcttcta gcatgtgata atgagtaaac taaaactatt 180
ttcagctttt ctcaattaac attttggtag tatacttcag agtgatgtta tctaagttta 240
agtagtttaa gtatgttaaa tgtggatctt ttacaccaca tcacagtgaa cacactgggg 300
agacgtgctt ttttggaaaa ctcaaaggtg ctagctccct gattcaaaga aatatttctc 360
atgtttgttc attctagttt atattttcat ttaaaatcct ttangttaag tttaagcttt 420
tt 422

<210> 64
<211> 221
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 12, 39, 45, 60, 63, 129, 130, 143, 144, 158
<223> n = A,T,C or G

<400> 64
agcttgatat cnaattcctg cagcccgggg gatccactng tccantgtgg tggaactcgn 60
cangactcag gacaatctcc agcatggcca gcttccctct cctcctcacc ctcctcactc 120
actgtgcann gtcctgggcc cannctgtgc tgactcancc accctcagcg tctgggaccc 180
ccggacagag ggtcaccatc tcttgttctg gaagcagctc c 221

<210> 65
<211> 520
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 55, 56, 180, 223, 235, 272, 289, 414
<223> n = A,T,C or G

<400> 65
tggaattccg cgacccggcg gcgggacagg cttgctgctt cctcctcctc ggccnnacca 60
ttccagacca aaattgaaaa aatggttgac ctcacccagg taatggatga tgaagtattc 120
atggcttttg catcctatgc aacaattatt ctttcaaaaa tgatgcttat gagtactgcn 180
actgcattct atagattgac aagaaaggtt tttgccaatc canaagactg tgtancattt 240
ggcaaaggag aaaatgccaa gaagtatctt cnaacagatg acagagtana acgtgtacgc 300
agagcccacc tgaatgacct tgaaaatatt attccatttc ttggaattgg cctcctgtat 360
tccttgagtg gtcccgaccc ctctacagcc atcctgcact tcagactatt tgtnggagca 420
cggatctacc acaccattgc atatttgaca ccccttcccc agccaaatag agctttgagt 480
ttttttgttg gatatggagt tactctttcc atggcttaca 520

<210> 66
<211> 392
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 379, 380
<223> n = A,T,C or G

<400> 66
aagctctgcc caaacaatct gtggatggaa aagcaccact tgctactgga gaggatgatg 60
atgatgaagt tccagatctt gtggagaatt ttgatgaggc ttccaagaat gaggcaaact 120
gaattgagtc aacttctgaa gataaaacct gaagaagtta ctgggagctg ctattttata 180
ttatgactgc tttttaagaa atttttgttt atggatctga taaaatctag atctctaata 240
tttttaagcc caagcccctt ggacactgca gctcttttca gtttttgctt atacacaatt 300
cattctttgc agctaattaa gccgaagaag cctgggaatc aagtttgaaa caaagattaa 360
taaagttctt tgcctagtnn aaaaaaaaaa aa 392

<210> 67
<211> 207
<212> DNA
<213> Homo sapiens

<400> 67

gaaatttaaa aactacaatg tgattaactc gagcctttag ttttcatcca tgtacatgga 60
tcacagtttg ctttgatctt cttcaatatg tgaatttggg ctcacagaat caaagcctat 120
gcttggttta atgcttgcaa tctgagctct tgaacaaata aaattaacta ttgtagtgtg 180
aaaaaaaaaa aaaaaaaggg cggccgg 207

<210> 68
<211> 373
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 366
<223> n = A,T,C or G

<400> 68
tacttcaaaa gaaaaataaa cataaaaaat aagttgctgg ttcctaacag gaaaaatttt 60
aacaattgta ctgagagaaa ctgcttacgt acacattgca gatcaaatat ttggagttaa 120
aatgttagtc tacatagatg ggtgattgta actttattgc cattaaaaga tttcaaattg 180
cattcatgct tctgtgtaca cataatgaaa aatgggcaaa taatgaagat ctctccttca 240
gtctgctctg tttaattctg ctgtctgctc ttctctaatg ctgcgtccct aattgtacac 300
agtttagtga tatctaggag tataaagttg tcgcccatca ataaaaatca caaagttggt 360
ttaaanaaaa aaa 373

<210> 69
<211> 367
<212> DNA
<213> Homo sapiens

<400> 69
tggaattcgc catcatggct gaccccgacc cccggtaccc tcgctcctcg atcgaggacg 60
acttcaacta tggcagcagc gtggcctccg ccaccgtgca catccgaatg gcctttctga 120
gaaaagtcta cagcattctt tctctgcagg ttctcttaac tacagtgact tcaacagttt 180
ttttatactt tgagtctgta cggacatttg tacatgagag tcctgcctta attttgctgt 240
ttgccctcgg atctctgggt ttgatttttg cgttgacttt aaacagacat aagtatcccc 300
ttaacctgta cctacttttt ggatttacgc tgttggaagc tctgactgtg gcagttgttg 360
ttacttt 367

<210> 70
<211> 568
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 18, 19, 522
<223> n = A,T,C or G

<400> 70
gtaactcctt catgcaanna actgaaaaga gccatgctgt ctagtcttga agtccctcat 60
ttaaacagag gtcaagcaat aggcgcctgg cagtgtcaag cctgaaacca agcaataccg 120
tcatgtttca gccaagccca gagccctaag attacaaaca actatggccg gaacctcctc 180
agctctccct ctgcagagtt ccctacccta agagaatgtt accacctgaa cagtcctcgg 240
tgaatctgag aggagaggat ggggtaaggc agaagcacca gctgtactac tagaagggag 300
cttttggtgg tagatcccct ggtgtctcca acctgactag gtggacagag ctcaaagagg 360
ccctcttacc gctagcgagg tgataggaca tctggcttgc cacaaaggtc tgttcgacca 420
gacatatcct agctaaggga tgtccaaaca tcagaatgtg aggccaacct tctatcagag 480
ttaaactttt gacaaaggga acaaatctca aactgatcca tnagtcatgt agctagctgt 540

agagcttgca acttaatagc agcagctg 568

<210> 71
<211> 483
<212> DNA
<213> Homo sapiens

<400> 71
tggaattccg ccaacatggg ccgcgttcgc accaaaaccg tgaagaaggc ggcccgggtc 60
atcatagaaa agtactacac gcgcctgggc aacgacttcc acacgaacaa gcgcgtgtgc 120
gaggagatcg ccattatccc cagcaaaaag ctccgcaaca agatagcagg ttatgtcacg 180
catctgatga agcgaattca gagaggccca gtaagaggta tctccatcaa gctgcaggag 240
gaggagagag aaaggagaga caattatgtt cctgaggtct cagccttgga tcaggagatt 300
attgaagtag atcctgacac taaggaaatg ctgaagcttt tggacttcgg cagtctgttc 360
aaccttcagg tcactcagcc tacagttggg atgaatttca aaacgcctcg gggacctgtt 420
tgaatttttt ctgtagtgct gtattatttt caataaatct gggacaacaa aaaaaaaaaa 480
aaa 483

<210> 72
<211> 452
<212> DNA
<213> Homo sapiens

<400> 72
tggaattcaa taactaaaag gtatgcaatc aaatctgctt tttaaagaat gctctttact 60
tcatggactt ccactgccat cctcccaagg ggcccaaatt ctttcagtgg ctacctacat 120
acaattccaa acacatacag gaaggtagaa atatctgaaa atgtatgtgt aagtattctt 180
atttaatgaa agactgtaca aagtagaagt cttagatgta tatatttcct atattgtttt 240
cagtgtacat ggaataacat gtaattaagt actatgtatc aatgagtaac aggaaaattt 300
taaaaataca gatagatata tgctctgcat gttacataag ataaatgtgc tgaatggttt 360
tcaaaataaa aatgaggtac tctcctggaa atattaagaa agactatcta aatgttgaaa 420
gaccaaaagg ttaataaagt aattataact aa 452

<210> 73
<211> 545
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 525
<223> n = A,T,C or G

<400> 73
ggccactgcg cagaccagac ttcgctcgta ctcgtgcgcc tcgcttcgct tttcctccgc 60
aaccatgtct gacaaacccg atatggctga gatcgagaaa ttcgataagt cgaaactgaa 120
gaagacagag acgcaagaga aaaatccact gccttccaaa gaaacgattg aacaggagaa 180
gcaagcaggc gaatcgtaat gaggcgtgcg ccgccaatat gcactgtaca ttccacaagc 240
attgccttct tattttactt cttttagctg tttaactttg taagatgcaa agaggttgga 300
tcaagtttaa atgactgtgc tgccccttc acatcaaaga actactgaca acgaaggccg 360
cgcctgcctt tcccatctgt ctatctatct ggctggcagg gaaggaaaga acttgcatgt 420
tggtgaagga agaagtgggg tggaagaagt ggggtgggac gacagtgaaa tctagagtaa 480
aaccaagctg gcccaaggtg tcctgcaggc tgtaatgcag tttantcaga gtgccatttt 540
ttttt 545

<210> 74
<211> 650
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<222> 564, 566, 606, 611, 634
<223> n = A,T,C or G

<400> 74
gattcactgg ggcattattt tgttagagga ccttaaaatt gtttattttt taaatgtgat 60
tcctttatgg cattagggta aagatgaagc aataattttt aaattgtgta tgtgcatatg 120
aagcacagac atgcatgtgt gtgtgtgtct gtgtgtgtgt gtccgtgtat gtgtgtgtgg 180
gttctaatgg taatttgcct cagtcatttt tttaatattt gcagtacttg atttaggatc 240
tgtggtgcag ggcaatgttt caaagtttag tcacagctta aaaacattca gtgtgacttt 300
aatattataa aatgatttcc catgccataa tttttctgtc tattaaatgg gacaagtgta 360
aagcatgcaa aagttagaga tctgttatat aacatttgtt ttgtgatttg aactcctagg 420
aaaaatatga tttcataaat gtaaaatgca cagaaatgca tgcaatactt ataagactta 480
aaaattgtgt tttacagatg gttttatttg tgcatatttt ttactactgc tttttcctaa 540
atgcatactg tatataaatt ctgngnattt gataaaatat ttccttccta cattatattt 600
ttagantatt ncagaaatat acatttatgt cttnatattg aaataaatat 650

<210> 75
<211> 506
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 172, 358, 400, 422
<223> n = A,T,C or G

<400> 75
atgctgcgcc tctccgaacg caacatgaag gtgctccttg ccgccgccct catcgcgggg 60
tccgtcttct tcctgctgct gccgggacct tctgcggccg atgagaagaa gaaggggccc 120
aaagtcaccg tcaaggtgta ttttgaccta cgaattggag atgaagatgt angccgggtg 180
atctttggtc tcttcggaaa gactgttcca aaaacagtgg ataattttgt ggccttagct 240
acaggagaga aaggatttgg ctacaaaaac agcaaattcc atcgtgtaat caaggacttc 300
atgatccagg gcggagactt caccagggga gatggcacag gaggaaagag catctacngt 360
gagcgcttcc ccgatgagaa cttcaaactg aagcactacn ggcctggctg ggtgagcatg 420
gncaacgcag gcaaagacac caacggctcc cagttcttca tcacgacagt caagacagcc 480
tggctagatg gcaagcatgt ggtgtt 506

<210> 76
<211> 543
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 370, 439, 445, 474, 518
<223> n = A,T,C or G

<400> 76
acgcagccgg ccaccgccga gacccagcac atcgccgacc aggtgaggtc ccagcttgaa 60
gagaaagaaa acaagaagtt ccctgtgttt aaggccgtgt cattcaagag ccaggtggtc 120
gcggggacaa actacttcat caaggtgcac gtcggcgacg aggacttcgt acacctgcga 180
gtgttccaat ctctccctca tgaaaacaag cccttgacct tatctaacta ccagaccaac 240
aaagccaagc atgatgagct gacctatttc tgatcctgac tttggacaag gcccttcagc 300
cagaagactg acaaagtcat cctccgtcta ccagagcgtg cacttgtgat cctaaaataa 360

gcttcatctn cgggctgtgc cccttggggt ggaaggggca ggattctgca gctgcttttg 420
catttctctt cctaaattnc attgngttga tttctttcct tcccaatagg tgancttaat 480
tactttcaga tatttttca aaaataagat atattttnta aaatcctaaa aaaaaaaaaa 540
aaa 543

<210> 77
<211> 535
<212> DNA
<213> Homo sapiens

<400> 77
gggaagcgtc tccgttgggt ccggccgctc tgcgggactc tgaggaaaag ctcgcaccag 60
gtggacgcgg atctgtcaac atgggtaaag gagaccccaa caagccgcgg ggcaaaatgt 120
cctcgtacgc cttcttcgtg cagacctgcc gggaagagca caagaagaaa cacccggact 180
cttccgtcaa tttcgcggaa ttctccaaga agtgttcgga gagatggaag accatgtctg 240
caaaggagaa gtcgaagttt gaagatatgg caaaaagtga caaagctcgc tatgacaggg 300
agatgaaaaa ttacgttcct cccaaaggtg ataagaaggg gaagaaaaag gaccccaatg 360
ctcctaaaag gccaccatct gccttcttcc tgttttgctc tgaacatcgc ccaaagatca 420
aaagtgaaca ccctggccta tccattgggg atactgcaaa gaaattgggt gaaatgtggt 480
ctgagcagtc agccaaagat aaacaaccat atgaacagaa agcagctaag ctaaa 535

<210> 78
<211> 595
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 491, 513
<223> n = A,T,C or G

<400> 78
tggaattcca taaagtacaa atgaagaaag tcaaaaaatt atttgctatg gcaggataag 60
aaagcctaaa attgagtttg tagaacttta ttaagtaaaa tccccttcgc tgaaattgct 120
tatttttggt gttggataga ggatagggag aatatttact aactaaatac cattcactac 180
tcatgcgtga gatgggtgta caaactcatc ctcttttaat ggcatttctc tttaaactat 240
gttcctaaca aaatgagatg ataggataga tcctggttac cactctttta ctgtgcacat 300
atgggctctg actggtttta atagtcacct tcatgattat agcaactaat gtttgaacaa 360
agctcaaagt atgcaatgct tcattattca agaatgaaaa atataatgtt gataatatat 420
attaagtgtg ccaaatcagt ttgactactc tctgttttag tgtttatgtt taaaagaaat 480
atatttttg ntattattag ataatatttt tgnatttctc tattttcata atcagtaaat 540
agtgtcatat aaactcattt atctcctctt catggcatct tcaatatgaa tctat 595

<210> 79
<211> 567
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 443, 448, 456
<223> n = A,T,C or G

<400> 79
agtcatactc cctggtgtag tgtattctct aaaagcttta aatgtctgca tgcagccagc 60
catcaaatag tgaatggtct ctctttggct ggaattacaa aactcagaga aatgtgtcat 120
caggagaaca tcataaccca tgaaggataa aagccccaaa tggtggtaac tgataatagc 180
actaatgctt taagatttgg tcacactctc acctaggtga gcgcattgag ccagtggtgc 240

taaatgctac atactccaac tgaaatgtta aggaagaaga tagatccaat taaaaaaaat 300
taaaaccaat ttaaaaaaaa aaagaacaca ggagattcca gtctacttga gttagcataa 360
tacagaagtc ccctctactt taacttttac aaaaaagtaa cctgaactaa tctgatgtta 420
accaatgtat ttatttctgt ggntctgntt ccttgntcca atttgacaaa acccactgtt 480
cttgtattgt attgcccagg gggagctatc actgtacttg tagagtggtg ctgctttaat 540
tcataaatca caaaataaaa gccaatt 567

<210> 80
<211> 155
<212> DNA
<213> Homo sapiens

<400> 80
gttccaatct ctccctcatg aaaacaagcc cttgacctta tctaactacc agaccaacaa 60
agccaagcat gatgagctga cctatttctg atcctgactt tggacaaggc ccttcagcca 120
gaagactgac aaaggcatcc tccgtctacc agagc 155

<210> 81
<211> 336
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 7, 110
<223> n = A,T,C or G

<400> 81
ctagttntgc cctcccgtca cccctgtttc tggcaccagg aatccccaac atgcactgat 60
gttgtgtttt taacatgtca atctgtccgt tcacatgtgt ggtacatggn gtttgtggcc 120
ttggctgaca tgaagctgtt gtgtgaggtt cgcttatcaa ctaatgattt agtgatcaaa 180
ttgtgcagta ctttgtgcat tctggatttt aaaagttttt tattatgcat tatatcaaat 240
ctaccactgt atgagtggaa attaagactt tatgtaggtt ttatatgttg taatatttct 300
tcaaataaat ctctcctata aaaaaaaaaa aaaagg 336

<210> 82
<211> 371
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 6, 24, 46, 48, 73, 81, 144, 194, 225, 227, 238, 247, 254,
279, 314, 340
<223> n = A,T,C or G

<400> 82
ctagtncagt gtggtggaat tcgnttgttg acccatctct gacagntnga gccgatatca 60
ctggaagata ttnaaaccgt ntctatgctt acgaacctgc agatacagct ctgttgcttg 120
acaacatgaa gaaagctctc aagntgctga agactgaatt gtaaagaaaa aaaatctcca 180
agcccttctg gctntcaggc cttgagactt gaaaccagaa gaagngngag aagactgnct 240
agtgtgnaag catngtgaac acactgatta ggttatggnt taatgttaca acaactattt 300
tttaagaaaa acangtttta gaaatttggt ttcaagtgtn catgtgtgaa aacaatattg 360
tatactacca t 371

<210> 83
<211> 386
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 37, 45, 57, 58, 95, 236, 377
<223> n = A,T,C or G

<400> 83
ctagtccagt gtggnggaat tcatctgacc atccatntcc aatgntctca tttaaannтt 60
acccagcatc attgtttata atcagaaact ctggnccttc tgtctggtgg cacttagagt 120
cttttgtgcc ataatgcagc agtatggagg gaggatttta tggagaaatg gggatagtct 180
tcatgaccac aaataaataa aggaaaacta agctgcattg tgggttttga aaaggntatt 240
atacttctta acaattcttt ttttcaggga cttttctagc tgtatgactg ttacttgacc 300
ttctttgaaa agcattccca aaatgctcta ttttagatag attaacatta accaacataa 360
tttttttag atcgagncag cataaa 386

<210> 84
<211> 381
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 229, 236, 318
<223> n = A,T,C or G

<400> 84
ctagtccagt gtggtggaat tcggccactg cgcagaccag acttcgctcg tactcgtgcg 60
cctcgcttcg cttttcctcc gcaaccatgt ctgacaaacc cgatatggct gagatcgaga 120
aattcgataa gtcgaaactg aagaagacag agacgcaaga gaaaaatcca ctgccttcca 180
aagaaacgat tgaacaggag aagcaagcag gcgaatcgta atgaggcgng cgccgncaaa 240
tatgcactgt acattccaca agcattgcct tcttatttta cttcttttag ctgtttaact 300
ttgtaagatg caaagagntt ggatcaagtt taaatgactg tgctgcccct ttcacatcaa 360
agaactactg acaacgaagg c 381

<210> 85
<211> 415
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 10, 15, 42, 73, 125
<223> n = A,T,C or G

<400> 85
ctagtccagn gtggnggaat tcctgaccag caccatggcg gntggcaaga acaagcgcct 60
tacgaaaggc ggnaaaaagg gagccaagaa gaaagtggtt gatccatttt ctaagaaaga 120
ttggnatgat gtgaaagcac ctgctatgtt caatataaga aatattggaa agacgctcgt 180
caccaggacc caaggaacca aaattgcatc tgatggtctc aagggtcgtg tgtttgaagt 240
gagtcttgct gatttgcaga atgatgaagt tgcatttaga aaattcaagc tgattactga 300
agatgttcag ggtaaaaact gcctgactaa cttccatggc atggatctta cccgtgacaa 360
aatgtgttcc atggtcaaaa aatggcagac aatgattgaa gctcacgttg atgtc 415

<210> 86
<211> 300
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 115
<223> n = A,T,C or G

<400> 86
ctagtgccat ttttgaaaaa agttggcttc aatcccaaaa aggacattca ctttatgccc 60
tgctcaggac ttactggagc aaatctcaaa gagcagtcgg atttctgtcc ttggnacatt 120
ggattaccgt ttattccata tctggataat ttgccgaact tcaatagatc agttgatgga 180
ccaatcaggc tgccaattgt ggataagtac aaggatatgg gcactgtggt cctgggaaag 240
ctggaatcag gatctatttg taaaggccag cagcttgtga tgatgccaaa caagcacaac 300

<210> 87
<211> 346
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 5, 12
<223> n = A,T,C or G

<400> 87
ctagnccagt gnggtggaat tccgcagcca tggctcgtgg tcccaagaag catctgaagc 60
gggtggcagc tccaaagcat tggatgctgg ataaattgac cggcgtgttt gctcctcgtc 120
catccaccgg tccccacaag ttgagagagt gtctcccccc catcattttc ctgaggaaca 180
gacttaagta tgccctgaca ggagatgaag taaagaagat ttgcatgcag cggttcatta 240
aaatcgatgg caaggtccga actgatataa cctaccctgc tggattcatg gatgtcatca 300
gcattgacaa gacgggagag aatttccgtc tgatctatga caccaa 346

<210> 88
<211> 238
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 143
<223> n = A,T,C or G

<400> 88
ctagtccagt gtggnggaat tccgagaaat tcgataagtc gaaactgaag aagacagaga 60
cgcaagagaa aaatccactg ccttccaaag aaacgattga acaggagaag caagcaggcg 120
aatcgtaatg aggcgtgcgc cgncaatatg cactgtacat tccacaagca ttgccttctt 180
attttacttc ttttagctgt ttaactttgt aagatgcaaa gaggttggat caagttta 238

<210> 89
<211> 316
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 194, 235, 273, 307, 309, 311
<223> n = A,T,C or G

<400> 89
ctagtccagt gtggtggaat tcggcgcgga gacgcttctg gaaggaacgc cgcgatggct 60
gcgcagggag agccccaggt ccagttcaaa cttgtattgg ttggtgatgg tggtactgga 120
aaaacgacct tcgtgaaacg tcatttgact ggtgaatttg agaagaagta tgtagccacc 180
ttgggtgttg aggntcatcc cctagtgttc cacaccaaca gaggacctat taagntcaat 240
gtatgggaca cagccggcca ggagaaattc ggnggactga gagatggcta ttatatccaa 300
gcccagngng ncatca 316

<210> 90
<211> 412
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 46, 68, 243, 305, 317, 364
<223> n = A,T,C or G

<400> 90
ctagttctgt ccccccagga gacctggttg tgtgtgtgtg agtggntgac cttcctccat 60
cccctggncc ttcccttccc ttcccgaggc acagagagac agggcaggat ccacgtgccc 120
attgtggagg cagagaaaag agaaagtgtt ttatatacgg gacttattta atatcccttt 180
ttaattagaa attaaaacag ttaatttaat taaagagtag ggtttttttt cagtattctt 240
ggntaatatt taatttcaac tatttatgag atgtatcttt tgctctctct tgctctctta 300
tttgnaccgg tttttgnata taaaattcat gtttccaatc tctctctccc tgatcgggga 360
cagncactag cttatcttga acagatattt aattttgcta acactcagct ct 412

<210> 91
<211> 271
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 257, 262
<223> n = A,T,C or G

<400> 91
ctagtccagt gtggnggaat tcgtctttct atctcttgta ctacactgaa ttcaccccca 60
ctgaaaaaga tgagtatgcc tgccgtgtga accatgtgac tttgtcacag cccaagatag 120
ttaagtggga tcgagacatg taagcagcat catggaggtt tgaagatgcc gcatttggat 180
tggatgaatt ccaaattctg cttgcttgct ttttaatatt gatatgctta tacacttaca 240
ctttatgcac aaaatgnagg gntataataa t 271

<210> 92
<211> 380
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 67, 149, 199, 208, 212, 342
<223> n = A,T,C or G

<400> 92
ctagtccagt gtggtggaat tcgcgcctta cgaaaggcgg caaaaaggga gccaagaaga 60
aagtggntga tccattttct aagaaagatt ggtatgatgt gaaagcacct gctatgttca 120
atataagaaa tattggaaag acgctcgtna ccaggaccca aggaaccaaa attgcatctg 180

atggtctcaa gggtcgtgng tttgaagnga gncttgctga tttgcagaat gatgaagttg 240
catttagaaa attcaagctg attactgaag atgttcaggg taaaaactgc ctgactaact 300
tccatggcat ggatcttacc cgtgacaaaa tgtggtccat gngcaaaaaa tggcagacaa 360
tgattgaagc tcacgttgat 380

<210> 93
<211> 354
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 285
<223> n = A,T,C or G

<400> 93
ctagtccagt gtggnaggaa ttcggagaat tcaagtgtga ccctcatgag gcaacgtgtt 60
atgatgatgg gaagacatac cacgtaggag aacagtggca gaaggaatat ctcggtgcca 120
tttgctcctg cacatgcttt ggaggccagc ggggctggcg ctgtgacaac tgccgcagac 180
ctgggggtga acccagtccc gaaggcacta ctggccagtc ctacaaccag tattctcaga 240
gataccatca gagaacaaac actaatgtta attgcccaat tgagngcttc atgcctttag 300
atgtacaggc tgacagagaa gattcccgag agtaaatcat ctttccaatc caga 354

<210> 94
<211> 247
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 244
<223> n = A,T,C or G

<400> 94
ctagtccagt gtggtggaat tccagcattc gggccgagat gtctcgctcc gtggccttag 60
ctgtgctcgc gctactctct ctttctggcc tggaggctat ccagcgtact ccaaagattc 120
aggtttactc acgtcatcca gcagagaatg gaaagtcaaa tttcctgaat tgctatgtgt 180
ctgggtttca tccatccgac attgaagttg acttactgaa gaatggagag agaattgaaa 240
aagngga 247

<210> 95
<211> 397
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 10, 15, 20, 42, 59, 69, 73, 125, 145, 240, 270
<223> n = A,T,C or G

<400> 95
ctagtccagn gtggnggaan tcctgaccag caccatggcg gntggcaaga acaagcgcnt 60
tacgaaagnc ggnaaaaagg gagccaagaa gaaagtggtt gatccatttt ctaagaaaga 120
ttggnatgat gtgaaagcac ctgcnatgtt caatataaga aatattggaa agacgctcgt 180
caccaggacc caaggaacca aaattgcatc tgatggtctc aagggtcgtg tgtttgaagn 240
gagtcttgct gatttgcaga atgatgaagn tgcatttaga aaattcaagc tgattactga 300
agatgttcag ggtaaaaact gcctgactaa cttccatggc atggatctta cccgtgacaa 360
aatgtgttcc atggtcaaaa aatggcagac aatgatt 397

<210> 96
<211> 287
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 92, 222, 237, 259
<223> n = A,T,C or G

<400> 96
ctagtccagt gtggtggaat tcggcgggtg aaaaagttga gaagccagat actaaagaga 60
agaaacccga agccaagaag gttgatgctg gnggcaaggt gaaaaagggt aacctcaaag 120
ctaaaaagcc caagaagggg aagccccatt gcagccgcaa ccctgtcctt gtcagaggaa 180
ttggcaggta ttcccgatct gccatgtatt ccagaaaggc cntgtacaag aggaagnact 240
cagccgctaa atccaaggnt gaaaagaaaa agaaggagaa ggttctc 287

<210> 97
<211> 387
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 32, 216, 219, 221, 302, 379
<223> n = A,T,C or G

<400> 97
ctagtccagt gtggtggaat tccgctcggc angttctccc aggagaaagc catgttcagt 60
tcgagcgcca agatcgtgaa gcccaatggc gagaagccgg acgagttcga gtccggcatc 120
tcccaggctc ttctggagct ggagatgaac tcggacctca aggctcagct cagggagctg 180
aatattacgg cagctaagga aattgaagtt ggtggnggnc nggaaagcta tcataatctt 240
tgttcccgtt cctcaactga aatctttcca gaaaatccaa gtccggctag tacgcgaatt 300
gnagaaaaag ttcagtggga agcatgtcgt ctttatcgct cagaggagaa ttctgcctaa 360
gccaactcga aaaagccgna caaaaaa 387

<210> 98
<211> 270
<212> DNA
<213> Homo sapiens

<400> 98
ctagtccagt gtggtggaat tcagcacctt caaagaaatc cccgtgactg tctatagacc 60
cacactaaca aaagtcaaaa ttgaaggtga acctgaattc agactgatta aagaaggtga 120
aacaataact gaagtgatcc atggagagcc aattattaaa aaatacacca aaatcattga 180
tggagtgcct gtggaaataa ctgaaaaaga gacacgagaa gaacgaatca ttacaggtcc 240
tgaaataaaa tacactagga tttctactgg 270

<210> 99
<211> 95
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 48, 76, 77, 83
<223> n = A,T,C or G

<400> 99
ctagtccagt gtggtggaat tcgcacagac agattgacct attggggngt ttcgcgagtg 60
tgagagggaa gcgccnnggc ctngtatttc tagac 95

<210> 100
<211> 312
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 10, 140, 207, 220, 227, 230, 247, 259
<223> n = A,T,C or G

<400> 100
ctagtccagn gtggtggaat tcgccgaaag gaaagaaggc caagggaaag aaggtggctc 60
cggccccagc tgtcgtgaag aagcaggagg ctaagaaagt ggtgaatccc ctgtttgaga 120
aaaggcctaa gaattttggn attggacagg acatccagcc caaaagagac ctcacccgct 180
ttgtgaaatg gccccgctat atcaggntgc agcggcagan agccatnctn tataagcggc 240
tgaaagngcc tcctgcgant aaccagttca cccaggccct ggaccgccaa acagctactc 300
agctgcttaa gc 312

<210> 101
<211> 395
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 232, 313
<223> n = A,T,C or G

<400> 101
ctagtccagt gtggtggaat tcactacgca gaccagactt cgctcgtact cgtgcgcctc 60
gcttcgcttt tcctccgcaa ccatgtctga caaacccgat atggctgaga tcgagaaatt 120
cgataagtcg aaactgaaga agacagagac gcaagagaaa aatccactgc cttccaaaga 180
aacgattgaa caggagaagc aagcaggcga atcgtaatga ggcgtgcgcc gncaatatgc 240
actgtacatt ccacaagcat tgccttctta ttttacttct tttagctgtt taactttgta 300
agatgcaaag agnttggatc aagtttaaat gactgtgctg ccccttttcac atcaaagaac 360
tactgacaac gaaggccgcg cctgcctttc ccatc 395

<210> 102
<211> 231
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 209
<223> n = A,T,C or G

<400> 102
ctagtgccta aatgtagtaa aggctgctta agttttgtat gtagttggat tttttggagt 60
ccgaaggtat ccatctgcag aaattgaggc ccaaattgaa tttggattca agtggattct 120
aaatactttg cttatcttga agagagaagc ttcataagga ataaacaagt tgaatagaga 180
aaacactgat tgataatagg cattttagng gccttttaa tgttttctgc t 231

<210> 103
<211> 399
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 324
<223> n = A,T,C or G

<400> 103
ctagtgtgtc tgatcagtga cttctacccg ggagctgtga cagtggcctg gaaggcagat 60
ggcagccccg tcaaggcggg agtggagacc accaaaccct ccaaacagag caacaacaag 120
tacgcggcca gcagctacct gagcctgacg cccgagcagt ggaagtccca cagaagctac 180
agctgccagg tcacgcatga agggagcacc gtggagaaga cagtggcccc tacagaatgt 240
tcataggttc ccaactctaa ccccacccac gggagcctgg agctgcagga tcccagggga 300
ggggtctctc tccccatccc aagncatcca gcccttctcc ctgcactcat gaaaccccaa 360
taaatatcct cattgacaac cagaaaaaaa aaaaaaaaa 399

<210> 104
<211> 370
<212> DNA
<213> Homo sapiens

<400> 104
ctagtccagt gtggtggaat tcggtggttt tcagtttagc tacggcaatc ctgaacttcc 60
tgaagatgtc cttgatgtgc agctggcatt ccttcgactt ctctccagcc gagcttccca 120
gaacatcaca tatcactgca aaaatagcat tgcatacatg gatcaggcca gtggaaatgt 180
aaagaaggcc ctgaagctga tggggtcaaa tgaaggtgaa ttcaaggctg aaggaaatag 240
caaattcacc tacacagttc tggaggatgg ttgcacgaaa cacactgggg aatggagcaa 300
aacagtcttt gaatatcgaa cacgcaaggc tgtgagacta cctattgtag atattgcacc 360
ctatgacatt 370

<210> 105
<211> 300
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 179
<223> n = A,T,C or G

<400> 105
ctagtccagt gtggtggaat tcgcggaggt gcaggtcctg gtgcttgatg gtcgaggcca 60
tctcctgggc cgcctggcgg ccatcgtggc taaacaggta ctgctgggcc ggaaggtggt 120
ggtcgtacgc tgtgaaggca tcaacatttc tggcaatttc tacagaaaca agttgaagna 180
cctggctttc ctccgcaagc ggatgaacac caacccttcc cgaggcccct accacttccg 240
ggcccccagc cgcatcttct ggcggaccgt gcgaggtatg ctgccccaca aaaccaagcg 300

<210> 106
<211> 349
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<222> 250
<223> n = A,T,C or G

<400> 106
ctagtccagt gtggtggaat tcaccgctcc aagcccagcc ctcagccatg gcatgccccc 60
tggatcaggc cattggcctc ctcgtggcca tcttccacaa gtactccggc agggagggtg 120
acaagcacac cctgagcaag aaggagctga aggagctgat ccagaaggag ctcaccattg 180
gctcgaagct gcaggatgct gaaattgcaa ggctgatgga agacttggac cggaacaagg 240
accaggaggn gaacttccag gagtatgtca ccttcctggg ggccttggct ttgatctaca 300
atgaagccct caagggctga aaataaatag ggaagatgga gacaccctc 349

<210> 107
<211> 298
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 214
<223> n = A,T,C or G

<400> 107
gcgagaagta cctgacttgg gcatcccggc aggagcccag ccagggcacc accaccttcg 60
ctgtgaccag catactgcgc gtggcagccg aggactggaa gaaggggggac accttctcct 120
gcatggtggg ccacgaggcc ctgccgctgg ccttcacaca gaagaccatc gaccgcttgg 180
cgggtaaacc cacccatgtc aatgtgtctg ttgncatggc ggaggtggac ggcacctgct 240
actgagccgc ccgcctgtcc ccacccctga ataaactcca tgctccccaa aaaaaaaa 298

<210> 108
<211> 135
<212> DNA
<213> Homo sapiens

<400> 108
ctagtccagt gtggtggaat tcggaccact gaagaaagac cgaattgcaa aggaagaagg 60
agcttaatgc caggaacaga ttttgcagtt ggtggggtct caataaaagt tattttccac 120
tgaaaaaaaa aaaaa 135

<210> 109
<211> 404
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 324
<223> n = A,T,C or G

<400> 109
ctagtgtgtc tgatcagtga cttctacccg ggagctgtga cagtggcctg gaaggcagat 60
ggcagccccg tcaaggcggg agtggagacc accaaaccct ccaaacagag caacaacaag 120
tacgcggcca gcagctacct gagcctgacg cccgagcagt ggaagtccca cagaagctac 180
agctgccagg tcacgcatga agggagcacc gtggagaaga cagtggcccc tacagaatgt 240
tcataggttc ccaactctaa ccccacccac gggagcctgg agctgcagga tcccagggga 300
ggggtctctc tccccatccc aagncatcca gcccttctcc ctgcactcat gaaaccccaa 360
taaatatcct cattgacaac cagaaaaaaa aaaaaaaaaa aggg 404

<210> 110

<211> 395
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 136, 244, 376
<223> n = A,T,C or G

<400> 110
ctagtgcttt acctttatta atgaactgtg acaggaagcc caaggcagtg ttcctcacca 60
ataacttcag agaagtcagt tggagaaaat gaagaaaaag gctggctgaa aatcactata 120
accatcagtt actggnttca gttgacaaaa tatataatgg tttactgctg tcattgtcca 180
tgcctacaga taatttattt tgtatttttg aataaaaaac atttgtacat tcctgatact 240
gggnacaaga gccatgtacc agtgtactgc tttcaactta aatcactgag gcatttttac 300
tactattctg ttaaaatcag gattttagtg cttgccacca ccagatgaga agttaagcag 360
cctttctgtg gagagngaga ataattgtgt acaaa 395

<210> 111
<211> 401
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 34, 164
<223> n = A,T,C or G

<400> 111
ctagtccagt gtggtggaat tccgaggctg cggngtctgc tgctattctc cgagcttcgc 60
aatgccgcct aaggacgaca agaagaagaa ggacgctgga aagtcggcca agaaagacaa 120
agacccagtg aacaaatccg ggggcaaggc caaaaagaag aagnggtcca aaggcaaagt 180
tcgggacaag ctcaataact tagtcttgtt tgacaaagct acctatgata aactctgtaa 240
ggaagttccc aactataaac ttataacccc agctgtggtc tctgagagac tgaagattcg 300
aggctccctg gccagggcag cccttcagga gctccttagt aaaggactta tcaaactggg 360
ttcaaagcac agagctcaag taatttacac cagaaatacc a 401

<210> 112
<211> 369
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 7, 81, 114, 261, 279, 280, 365
<223> n = A,T,C or G

<400> 112
ctagtcnagt gtggtggaat tcggctggta agcaggccgt ttcagcatca ggcaagtggc 60
tggatggtat tcgaaaatgg nattacaatg ctgcaggatt caataaactg gggntaatgc 120
gagatgatac aatatacgag gatgaagatg taaaagaagc cataagaaga cttcctgaga 180
acctttataa tgacaggatg tttcgcatta agagggcact ggacctgaac ttgaagcatc 240
agatcttgcc taaagagcag nggaccaaat atgaagagnn aaatttctac cttgaaccgt 300
atctgaaaga ggttattcgg gaaagaaaag aaagagaaga atgggcaaag aagtaatcat 360
gtagntgaa 369

<210> 113
<211> 56

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 5, 49, 51
<223> n = A,T,C or G

<400> 113
ctagntatta atagtaatca attacggggt cattagttca tagcccatnt ntggag 56

<210> 114
<211> 361
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 358
<223> n = A,T,C or G

<400> 114
ctagtccagt gtggtggaat tcattctcag caatcagact gtcgacattc cagaaaatgt 60
cgacattact ctgaagggac gcacagttat cgtgaagggc cccagaggaa ccctgcggag 120
ggacttcaat cacatcaatg tagaactcag ccttcttgga aagaaaaaaa agaggctccg 180
ggttgacaaa tggtggggta acagaaagga actggctacc gttcggacta tttgtagtca 240
tgtacagaac atgatcaagg gtgttacact gggcttccgt tacaagatga ggtctgtgta 300
tgctcacttc cccatcaacg ttgttatcca ggagaatggg tctcttgttg aaatccgnaa 360
t 361

<210> 115
<211> 310
<212> DNA
<213> Homo sapiens

<400> 115
ctagtccagt gtggtggaat tcatgacaac aaatggtgta attcatgttg tagataaact 60
cctctatcca gcagacacac ctgttggaaa tgatcaactg ctggaaatac ttaataaatt 120
aatcaaatac atccaaatta agtttgttcg tggtagcacc ttcaaagaaa tccccgtgac 180
tgtctataag ccaattatta aaaaatacac caaaatcatt gatggagtgc ctgtggaaat 240
aactgaaaaa gagacacgag aagaacgaat cattacaggt cctgaaataa aatacactag 300
gatttctact 310

<210> 116
<211> 278
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 11, 20, 30, 106, 129, 148, 214
<223> n = A,T,C or G

<400> 116
caaagtctcg nttctgccgn ggtgtccctn atgccaagat tcgcattttt gacctggggc 60
ggaaaaaggc aaaagtggat gagtttccgc tttgtggcca catggngtca gatgaatatg 120
agcagctgnc ctctgaagcc ctggaggntg cccgaatttg tgccaataag tacatggtaa 180
aaagttgtgg caaagatggc ttccatatcc gggngcggct ccacccccttc cacgtcatcc 240

gcatcaacaa gatgttgtcc tgtgctgggc tgacaggc 278

<210> 117
<211> 233
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 88, 211
<223> n = A,T,C or G

<400> 117
tcaacatgaa ggctctcatt gttctggggc ttgtcctcct ttctgttacg gtccagggca 60
aggtctttga aaggtgtgag ttggccanaa ctctgaaaag attgggaatg gatggctaca 120
ggggaatcag cctagcaaac tggatgtgtt tggccaaatg ggagagtggt tacaacacac 180
gagctacaaa ctacaatgct ggagacagaa ncactgatta tgggatattt cag 233

<210> 118
<211> 552
<212> DNA
<213> Homo sapiens

<400> 118
ctagtccagt gtggtggaat tctaagatgg aagcgttttt ggggtcgcgg tccggacttt 60
gggcgggggg tccggcccca ggacagtttt accgcattcc gtccactccc gattccttca 120
tggatccggc gtctgcactt tacagaggtc caatcacgcg gacccagaac cccatggtga 180
ccgggacctc agtcctcggc gttaagttcg agggcggagt ggtgattgcc gcagacatgc 240
tgggatccta cggctccttg gctcgtttcc gcaacatctc tcgcattatg cgagtcaaca 300
acagtaccat gctgggtgcc tctggcgact acgctgattt ccagtatttg aagcaagttc 360
tcggccagat ggtgattgat gaggagcttc tgggagatgg acacagctat agtcctagag 420
ctattcattc atggctgacc agggccatgt acagccggcg ctcgaagatg aacccttttgt 480
ggaacaccat ggtcatcgga ggctatgctg atggagagag cttcctcggt tatgtggaca 540
tgcttggtgt ag 552

<210> 119
<211> 465
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 14, 17, 18, 340, 356, 359, 375, 448, 449, 450
<223> n = A,T,C or G

<400> 119
ctagtccagt gtgntgnnat tcgtaggagg gatttcggcc tgagagcggg ccgaggagat 60
tggcgacggt gtcgcccgtg ttttcgttgg cgggtgcctg ggctggtggg aacagccgcc 120
cgaaggaagc accatgattt cggccgcgca gttgttggat gagttaatgg gccgggaccg 180
aaacctagcc ccggacgaga agcgcagcaa cgtgcggtgg gaccacgaga gcgtttgtaa 240
atattatctc tgtggttttt gtcctgcgga attgttcaca aatacacgtt ctgatcttgg 300
tccgtgtgaa aaaattcatg atgaaaatct acgaaaacan tatgagaaga gctctngtnt 360
catgaaagtt ggctntgaga gagatttttt gcgatactta cagagcttac ttgcagaagt 420
agaacgtagg atcagacgag gccatgcnnn gtttggcatt atctc 465

<210> 120
<211> 50
<212> DNA

<213> Homo sapiens

<400> 120
ctagcgttta aacttaagct tggtaccgag ctcggatctc gagtctagag 50

<210> 121
<211> 281
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 162, 215, 229
<223> n = A,T,C or G

<400> 121
aattccttgg ctcctgtgga ggcctgctgg gaacgggact tctaaaagga actatgtctg 60
gaaggctgtg gtccaaggcc atttttgctg gctataagcg gggtctccgg aaccaaaggg 120
agcacacagc tcttcttaaa attgaaggtg tttacgcccg anatgaaaca gaattctatt 180
tgggcaagag atgcgcttat gtatataaag caaanaacaa cacagtcant cctggcggca 240
aaccaaacaa aaccagagtc atctggggaa aagtaactcg g 281

<210> 122
<211> 221
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 11, 121, 147, 152
<223> n = A,T,C or G

<400> 122
caagactact ntaccctgca acattgaact cccaagagca aatccacatt cctcttgagt 60
tctgcagctt ctgtgtaaat agggcagctg tcgtctatgc cgtagaatca catgatctga 120
ngaccattca tggaagctgc taaatancct antctgggga gtcttccata aagttttgca 180
tggagcaaac aaacaggatt aaactaggtt tggttccttc a 221

<210> 123
<211> 557
<212> DNA
<213> Homo sapiens

<400> 123
ctagtccagt gtggtggaat tcggcctaca cgccgccgct tgtgctgcag ccatgtctct 60
agtgatccct gaaaagttcc agcatatttt gcgagtactc aacaccaaca tcgatgggcg 120
gcggaaaata gcctttgcca tcactgccat taagggtgtg ggccgaagat atgctcatgt 180
ggtgttgagg aaagcagaca ttgacctcac caagagggcg ggagaactca ctgaggatga 240
ggtggaacgt gtgatcacca ttatgcagaa tccacgccag tacaagatcc cagactggtt 300
cttgaacaga cagaaggatg taaaggatgg aaaatacagc caggtcctag ccaatggtct 360
ggacaacaag ctccgtgaag acctggagcg actgaagaag attcgggccc atagagggct 420
gcgtcacttc tggggccttc gtgtccgagg ccagcacacc aagaccactg gccgccgtgg 480
ccgcaccgtg ggtgtgtcca agaagaaata agtctgtagg ccttgtctgt taataaatag 540
tttatatacc taaaaaa 557

<210> 124
<211> 532
<212> DNA

<213> Homo sapiens

<400> 124
ctagttttta agaagaaatt ttttttggcc tatgaaattg ttaaacctgg aacatgacat 60
tgttaatcat ataataatga ttcttaaatg ctgtatggtt tattatttaa atgggtaaag 120
ccatttacat aatatagaaa gatatgcata tatctagaag gtatgtggca tttatttgga 180
taaaattctc aattcagaga aatcatctga tgtttctata gtcactttgc cagctcaaaa 240
gaaaacaata ccctatgtag ttgtggaagt ttatgctaat attgtgtaac tgatattaaa 300
cctaaatgtt ctgcctaccc tgttggtata aagatatttt gagcagactg taaacaagaa 360
aaaaaaaatc atgcattctt agcaaaattg cctagtatgt taatttgctc aaaatacaat 420
gtttgatttt atgcactttg tcgctattaa catccttttt ttcatgtaga tttcaataat 480
tgagtaattt tagaagcatt attttaggaa tatatagttg tcacagtaaa ta 532

<210> 125
<211> 558
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 409, 554
<223> n = A,T,C or G

<400> 125
ctagtccagt gtggtggaat tcgcaagttc tcccaggaga aagccatgtt cagttcgagc 60
gccaagatcg tgaagcccaa tggcgagaag ccggacgagt tcgagtccgg catctcccag 120
gctcttctgg agctggagat gaactcggac ctcaaggctc agctcaggga gctgaatatt 180
acggcagcta aggaaattga agttggtggt ggtcggaaag ctatcataat ctttgttccc 240
gttcctcaac tgaaatcttt ccagaaaatc caagtccggc tagtacgcga attggagaaa 300
aagttcagtg ggaagcatgt cgtctttatc gctcagagga gaattctgcc taagccaact 360
cgaaaaagcc gtacaaaaaa taagcaaaag cgtcccagga gccgtactnt gacagctgtg 420
cacgatgcca tccttgagga cttggtcttc ccaagcgaaa ttgtgggcaa gagaatccgc 480
gtcaaactag atggcagccg gctcataaag gttcatttgg acaaagcaca gcagaacaat 540
gtggaacaca aggntgaa 558

<210> 126
<211> 575
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 558, 559, 560
<223> n = A,T,C or G

<400> 126
ctagtccagt gtggtggaat tcgcggcagc catcaggtaa gccaagatgg gtgcatacaa 60
gtacatccag gagctatgga gaaagaagca gtctgatgtc atgcgctttc ttctgagggt 120
ccgctgctgg cagtaccgcc agctctctgc tctccacagg gctccccgcc ccacccggcc 180
tgataaagcg cgccgactgg gctacaaggc caagcaaggt tacgttatat ataggattcg 240
tgttcgccgt ggtggccgaa aacgcccagt tcctaagggt gcaacttacg gcaagcctgt 300
ccatcatggt gttaaccagc taaagtttgc tcgaagcctt cagtccgttg cagaggagcg 360
agctggacgc cactgtgggg ctctgagagt cctgaattct tactgggttg gtgaagattc 420
cacatacaaa ttttttgagg ttatcctcat tgatccattc cataaagcta tcagaagaaa 480
tcctgacacc cagtggatca ccaaaccagt ccacaagcac agggagatgc gtgggctgac 540
atctgcaggc cgaaagannn gtggccttgg aaagg 575

<210> 127

<211> 614
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 554, 587
<223> n = A,T,C or G

<400> 127
ctagtccagt gtggtggaat tcgggtactc aacactgagc agatctgttc tttgagctaa 60
aaaccatgtg ctgtaccaag agtttgctcc tggctgcttt gatgtcagtg ctgctactcc 120
acctctgcgg cgaatcagaa gcaagcaact ttgactgctg tcttggatac acagaccgta 180
ttcttcatcc taaatttatt gtgggcttca cacggcagct ggccaatgaa ggctgtgaca 240
tcaatgctat catctttcac acaaagaaaa agttgtctgt gtgcgcaaat ccaaaacaga 300
cttgggtgaa atatattgtg cgtctcctca gtaaaaaagt caagaacatg taaaaactgt 360
ggcttttctg gaatggaatt ggacatagcc caagaacaga aagaaccttg ctggggttgg 420
aggtttcact tgcacatcat ggagggttta gtgcttatct aatttgtgcc tcactggact 480
tgtccaatta atgaagttga ttcatattgc atcatagttt gctttgttta agcatcacat 540
taaagttaaa ctgnatttta tgttatttat agctgtaggt tttctgngtt tagctattta 600
atactaattt tcca 614

<210> 128
<211> 420
<212> DNA
<213> Homo sapiens

<400> 128
ctagtttaag gagactggcc gaagctctgc ccaaacaatc tgtggatgga aaagcaccac 60
ttgctactgg agaggatgat gatgatgaag ttccagatct tgtggagaat tttgatgagg 120
cttccaagaa tgaggcaaac tgaattgagt caacttctga agataaaacc tgaagaagtt 180
actgggagct gctattttat attatgactg ctttttaaga aatttttgtt tatggatctg 240
ataaaatcta gatctctaat atttttaagc ccaagcccct tggacactgc agctcttttc 300
agtttttgct tatacacaat tcattctttg cagctaatta agccgaagaa gcctgggaat 360
caagtttgaa acaaagatta ataaagttct ttgcctagta aaaaaaaaaa aaaaaagggc 420

<210> 129
<211> 416
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 10, 14, 15, 27, 82, 219, 239, 268, 289, 290, 307, 344, 382,
389, 394, 407
<223> n = A,T,C or G

<400> 129
ctagtccagn gtgnntggaa ttcgtcnaag cgaggacgtg gtgggtcctc tggtgcgaaa 60
ttccggattt ccttgggtct tncggtagga gctgtaatca attgtgctga caacacagga 120
gccaaaaacc tgtatatcat ctccgtgaag gggatcaagg gacggctgaa cagacttccc 180
gctgctggtg tgggtgacat ggtgatggcc acagtcaana aaggcaaacc agagctcana 240
aaaaaggtac atccagcagt ggtcattnga caacgaaagt cataccgtnn aaaagatggc 300
gtgtttnttt attttgaaga taatgcagga gtcatagtga acantaaagg cgagatgaaa 360
ggttctgcca ttacaggacc angtagcana ggantgtgca gacttgnggc ccccgg 416

<210> 130

<211> 623
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 560, 593
<223> n = A,T,C or G

<400> 130
ctagtccagt gtggtggaat tcagaactgg gtactcaaca ctgagcagat ctgttctttg 60
agctaaaaac catgtgctgt accaagagtt tgctcctggc tgctttgatg tcagtgctgc 120
tactccacct ctgcggcgaa tcagaagcaa gcaactttga ctgctgtctt ggatacacag 180
accgtattct tcatcctaaa tttattgtgg gcttcacacg gcagctggcc aatgaaggct 240
gtgacatcaa tgctatcatc tttcacacaa agaaaaagtt gtctgtgtgc gcaaatccaa 300
aacagacttg ggtgaaatat attgtgcgtc tcctcagtaa aaaagtcaag aacatgtaaa 360
aactgtggct tttctggaat ggaattggac atagcccaag aacagaaaga accttgctgg 420
ggttggaggt ttcacttgca catcatggag ggtttagtgc ttatctaatt tgtgcctcac 480
tggacttgtc caattaatga agttgattca tattgcatca tagtttgctt tgtttaagca 540
tcacattaaa gttaaactgn attttatgtt atttatagct gtaggttttc tgngtttagc 600
tatttaatac taattttcca taa 623

<210> 131
<211> 439
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 14, 15, 17, 29, 305, 424
<223> n = A,T,C or G

<400> 131
ctagtccagt gtgnngnaat tccttgacna ggctgcggtg tctgctgcta ttctccgagc 60
ttcgcaatgc cgcctaagga cgacaagaag aagaaggacg ctggaaagtc ggccaagaaa 120
gacaaagacc cagtgaacaa atccgggggc aaggccaaaa agaagaagtg gtccaaaggc 180
aaagttcggg acaagctcaa taacttagtc ttgtttgaca aagctaccta tgataaactc 240
tgtaaggaag ttcccaacta taaacttata accccagctg tggtctctga gagactgaag 300
attcnaggct ccctggccag ggcagccctt caggagctcc ttagtaaagg acttatcaaa 360
ctggtttcaa agcacagagc tcaagtaatt tacaccagaa ataccaaggg tggagatgct 420
ccanctgctg gtgaagatg 439

<210> 132
<211> 619
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 557
<223> n = A,T,C or G

<400> 132
ctagtccagt gtggtggaat tcgacagcat tcgggccgag atgtctcgct ccgtggcctt 60
agctgtgctc gcgctactct ctctttctgg cctggaggct atccagcgta ctccaaagat 120
tcaggtttac tcacgtcatc cagcagagaa tggaaagtca aatttcctga attgctatgt 180
gtctgggttt catccatccg acattgaagt tgacttactg aagaatggag agagaattga 240
aaaagtggag cattcagact tgtctttcag caaggactgg tctttctatc tcttgtacta 300

cactgaattc acccccactg aaaaagatga gtatgcctgc cgtgtgaacc atgtgacttt 360
gtcacagccc aagatagtta agtgggatcg agacatgtaa gcagcatcat ggaggtttga 420
agatgccgca tttggattgg atgaattcca aattctgctt gcttgctttt taatattgat 480
atgcttatac acttacactt tatgcacaaa atgtagggtt ataataatgt taacatggac 540
atgatcttct ttataanttc tactttgagt gctgtctcca tgtttgatgt atctgagcag 600
gttgctccac aggtagctc 619

<210> 133
<211> 583
<212> DNA
<213> Homo sapiens

<400> 133
ctagtccagt gtggtggaat tcaagaggag gaagctgtta ccatagagat gaatgaacca 60
gttcaactaa cttttgcact gaggtacctg aacttcttta caaaagccac tccactctct 120
tcaacggtga cactcagtat gtctgcagat gtacccctg ttgtagagta taaaattgcg 180
gatatgggac acttaaaata ctacttggct cccaagatcg aggatgaaga aggatcttag 240
gcattcttaa aattcaagaa aataaaacta agctctttga gaactgcttc taagatgcca 300
gcatatactg aagtcttttc tgtcaccaaa tttgtacctc taagtacata tgtagatatt 360
gttttctgta aataacctat ttttttctct attctctgca atttgtttaa agaataaagt 420
ccaaagtcag atctggtcta gttaacctag aagtattttt gtctcttaga aatacttgtg 480
atttttataa tacaaaaggg tcttgactct aaatgcagtt ttaagaattg tttttgaatt 540
taaataaagt tacttgaatt tcaaaaaaaa aaaaaaaaag ggc 583

<210> 134
<211> 481
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 17, 373
<223> n = A,T,C or G

<400> 134
ctagtccagt gtggtgnaat tcgcgccgct ccggctgcac cgcgctcgct ccgagtttca 60
ggctcgtgct aagctagcgc cgtcgtcgtc tcccttcagt cgccatcatg attatctacc 120
gggacctcat cagccacgat gagatgttct ccgacatcta caagatccgg gagatcgcgg 180
acgggttgtg cctggaggtg gaggggaaga tggtcagtag gacagaaggt aacattgatg 240
actcgctcat tggtggaaat gcctccgctg aaggccccga gggcgaaggt accgaaagca 300
cagtaatcac tggtgtcgat attgtcatga accatcacct gcaggaaaca agtttcacaa 360
aagaagccta canagaagta catcaaagat tacatgaaat caatcaaagg gaaacttgaa 420
gaacagagac cagaaagagt aaaacctttt atgacagggg ctgcagaaca aatcaagcac 480
a 481

<210> 135
<211> 383
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 364, 365
<223> n = A,T,C or G

<400> 135
tggaattcgc cgcagaagcg agatgacgaa gggaacgtca tcgtttggaa agcgtcgcaa 60
taagacgcac acgttgtgcc gccgctgtgg ctctaaggcc taccaccttc agaagtcgac 120

ctgtggcaaa tgtggctacc ctgccaagcg caagagaaag tataactgga gtgccaaggc 180
taaaagacga ataccaccg gaactggtcg aatgaggcac ctaaaaattg tataccgcag 240
attcaggcat ggattccgtg aaggaacaac acctaaaccc aagagggcag ctgttgcagc 300
atccagttca tcttaagaat gtcaacgatt agtcatgcaa taaatgttct ggttttaaaa 360
aatnnaaaaa aaaaaaaaag ggc 383

<210> 136
<211> 629
<212> DNA
<213> Homo sapiens

<400> 136
ctagtccagt gtggtggaat tctgacaaca gcctcaagat catcagcaat gcctcctgca 60
ccaccaactg cttagcaccc ctggccaagg tcatccatga caactttggt atcgtggaag 120
gactcatgac cacagtccat gccatcactg ccacccagaa gactgtggat ggcccctccg 180
ggaaactgtg gcgtgatggc cgcggggctc tccagaacat catccctgcc tctactggcg 240
ctgccaaggc tgtgggcaag gtcatccctg agctgaacgg gaagctcact ggcatggcct 300
tccgtgtccc cactgccaac gtgtcagtgg tggacctgac ctgccgtcta gaaaaacctg 360
ccaaatatga tgacatcaag aaggtggtga agcaggcgtc ggagggcccc ctcaagggca 420
tcctgggcta cactgagcac caggtggtct cctctgactt caacagcgac acccactcct 480
ccacctttga cgctggggct ggcattgccc tcaacgacca ctttgtcaag ctcatttcct 540
ggtatgacaa cgaatttggc tacagcaaca gggtggtgga cctcatggcc cacatggcct 600
ccaaggagta agacccctgg accaccagc 629

<210> 137
<211> 227
<212> DNA
<213> Homo sapiens

<400> 137
ctagtcttga acaaactgtc atacgtatgg gacctacact taatctatat gctttacact 60
agctttctgc atttaatagg ttagaatgta aattaaagtg tagcaatagc aacaaaatat 120
ttattctact gtaaatgaca aaagaaaaag aaaaattgag ccttgggacg tgcccatttt 180
tactgtaaat tatgattccg taactgactt gtagtaagca gtgtttc 227

<210> 138
<211> 572
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 247
<223> n = A,T,C or G

<400> 138
ctagttatct tttaaaaggc tcagcaacac aactcttgaa atgcttatca ggataatggt 60
agctatagct ggccatttag aggaattcta ggacagtggg agctgtgtta ctagcactat 120
ataattccgg tcagtgctga caaataacat ttaacaagta ttgcagtaat catcacttac 180
aggtaccatt tatttcaaaa caacttttt agtctgctcc aaagttaaaa taattaacta 240
gctaagnatt attattcgac tggtctaaaa actattgtta tctttttttt ttcctttca 300
ctgttatggc cttttcacat ttctaaatcc catcttgata tactatgaat actctagaat 360
gatgtaaagc agataggaat gtatgtgtac atatttattg catacttgca catcaaatcg 420
atgtacatag tttaacacgt ggtccttttg tgaaacctag aactcagagg attgcttttt 480
ttctttcagc ctattttgag ttaacttcag tgctttctta gggaaatgac agggcaaagc 540
aatttttctg ttggctttgg gctgtatttg tg 572

<210> 139

<211> 576
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 235, 236, 240, 247, 445, 448, 495
<223> n = A,T,C or G

<400> 139
ctagtagtca tactccctgg tgtagtgtat tctctaaaag ctttaaatgt ctgcatgcag 60
ccagccatca aatagtgaat ggtctctctt tggctggaat tacaaaactc agagaaatgt 120
gtcatcagga gaacatcata acccatgaag gataaaagcc ccaaatggtg gtaactgata 180
atagcactaa tgctttaaga tttggtcaca ctctcaccta ggtgagcgca ttganncagn 240
ggtgctnaat gctacatact ccaactgaaa tgttaaggaa gaagatagat ccaattaaaa 300
aaaattaaaa ccaatttaaa aaaaaaaaga acacaggaga ttccagtcta cttgagttag 360
cataatacag aagtcccctc tactttaact tttacaaaaa agtaacctga actaatctga 420
tgttaaccaa tgtatttatt tctgnggntc tgtttccttg ttccaatttg acaaaaccca 480
ctgttcttgt attgnattgc ccaggggggag ctatcactgt acttgtagag tggtgctgct 540
ttaattcata aatcacaaat aaaagccaat tagctc 576

<210> 140
<211> 429
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 9, 25, 148, 192, 235, 267, 288, 293, 298, 326, 332, 333,
376, 394, 418
<223> n = A,T,C or G

<400> 140
aattcgcana ccagacttcg ctcgnactcg tgcgcctcgc ttcgcttttc ctccgcaacc 60
atgtctgaca aacccgatat ggctgagatc gagaaattcg ataagtcgaa actgaagaag 120
acagagacgc aagagaaaaa tccactgnct tccaaagaaa cgattgaaca ggagaagcaa 180
gcaggcgaat cntaatgagg cgtgcgccgc caatatgcac tgtacattcc acaancattg 240
ccttcttatt ttacttcttt tagctgntta actttgtaag atgcaaanag gtnggatnaa 300
gtttaaatga ctgtgctgcc cctttnacat cnnagaacta ctgacaacga aggccgcgcc 360
tgcctttccc atctgnctat ctatctggct ggcngggaag gaaagaactt gcatgttngt 420
gaaggaaga 429

<210> 141
<211> 624
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 178, 268, 498, 615, 617
<223> n = A,T,C or G

<400> 141
ctagtccagt gtggtggaat tccagcattc gggccgagat gtctcgctcc gtggccttag 60
ctgtgctcgc gctactctct ctttctggcc tggaggctat ccagcgtact ccaaagattc 120
aggtttactc acgtcatcca gcagagaatg gaaagtcaaa tttcctgaat tgctatgngt 180
ctgggtttca tccatccgac attgaagttg acttactgaa gaatggagag agaattgaaa 240
aagtggagca ttcagacttg tctttcanca aggactggtc tttctatctc ttgtactaca 300

ctgaattcac ccccactgaa aaagatgagt atgcctgccg tgtgaaccat gtgactttgt 360
cacagcccaa gatagttaag tgggatcgag acatgtaagc agcatcatgg aggtttgaag 420
atgccgcatt tggattggat gaattccaaa ttctgcttgc ttgcttttta atattgatat 480
gcttatacac ttacactnta tgcacaaaat gtagggttat aataatgtta acatggacat 540
gatcttcttt ataattctac tttgagtgct gtctccatgt ttgatgtatc tgagcaggtt 600
gctccacagg tagcntntag gagg 624

<210> 142
<211> 626
<212> DNA
<213> Homo sapiens

<400> 142
ctagttttaa gatcagagtt cactttcttt ggactctgcc tatattttct tacctgaact 60
tttgcaagtt ttcaggtaaa cctcagctca ggactgctat ttagctcctc ttaagaagat 120
taaaagagaa aaaaaaaggc cttttaaaa atagtataca cttattttaa gtgaaaagca 180
gagaatttta tttatagcta attttagcta tctgtaacca agatggatgc aaagaggcta 240
gtgcctcaga gagaactgta cggggtttgt gactggaaaa agttacgttc ccattctaat 300
taatgccctt tcttatttaa aaacaaaacc aaatgatatc taagtagttc tcagcaataa 360
taataatgac gataatactt cttttccaca tctcattgtc actgacattt aatggtactg 420
tatattactt aatttattga agattattat ttatgtctta ttaggacact atggttataa 480
actgtgttta agcctacaat cattgatttt tttttgttat gtcacaatca gtatattttc 540
tttggggtta cctctctgaa tattatgtaa acaatccaaa gaaatgattg tattaagatt 600
tgtgaataaa tttttagaaa tctgat 626

<210> 143
<211> 554
<212> DNA
<213> Homo sapiens

<400> 143
ctagttttta agaagaaatt ttttttggcc tatgaaattg ttaaacctgg aacatgacat 60
tgttaatcat ataataatga ttcttaaatg ctgtatggtt tattatttaa atgggtaaag 120
ccatttacat aatatagaaa gatatgcata tatctagaag gtatgtggca tttatttgga 180
taaaattctc aattcagaga aatcatctga tgtttctata gtcactttgc cagctcaaaa 240
gaaaacaata ccctatgtag ttgtggaagt ttatgctaat attgtgtaac tgatattaaa 300
cctaaatgtt ctgcctaccc tgttggtata aagatatttt gagcagactg taaacaagaa 360
aaaaaaaatc atgcattctt agcaaaattg cctagtatgt taatttgctc aaaatacaat 420
gtttgatttt atgcactttg tcgctattaa catccttttt ttcatgtaga tttcaataat 480
tgagtaattt tagaagcatt attttaggaa tatatagttg tcacagtaaa tatcttgttt 540
tttctatgta catt 554

<210> 144
<211> 345
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 94, 99, 120, 197, 208, 215, 258, 270, 309, 311, 339
<223> n = A,T,C or G

<400> 144
ctagttttta agaanaaatt ttttttggcc tatgaaattg ttaaacctgg aacatgacat 60
tgttaatcat ataataatga ttcttaaatg ctgnatggnt tattatttaa atgggtaaan 120
ccatttacat aatatagaaa gatatgcata tatctagaag gtatgtggca tttatttgga 180
taaaattctc aattcanaga aatcatcnga tgttnctata gtcactttgc cagctcaaaa 240
gaaaacaata ccctatgnag ttgtggaagn ttatgctaat attgtgtaac tgatattaaa 300

cctaaatgnt ntgcctaccc tgttggtata aagatattnt gagca 345

<210> 145
<211> 477
<212> DNA
<213> Homo sapiens

<400> 145
ctagttttta agaagaaatt ttttttggcc tatgaaattg ttaaacctgg aacatgacat 60
tgttaatcat ataataatga ttcttaaatg ctgtatggtt tattatttaa atgggtaaag 120
ccatttacat aatatagaaa gatatgcata tatctagaag gtatgtggca tttatttgga 180
taaaattctc aattcagaga aatcatctga tgtttctata gtcactttgc cagctcaaaa 240
gaaaacaata ccctatgtag ttgtggaagt ttatgctaat attgtgtaac tgatattaaa 300
cctaaatgtt ctgcctaccc tgttggtata aagatatttt gagcagactg taaacaagaa 360
aaaaaaaatc atgcattctt agcaaaattg cctagtatgt taatttgctc aaaatacaat 420
gtttgatttt atgcactttg tcgctattaa catccttttt ttcatgtagg atttcaa 477

<210> 146
<211> 512
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 463, 485, 496
<223> n = A,T,C or G

<400> 146
ctagtccagt gtggtggaat tcagataagt gtccatagcc tgtttctgtc attaatgagc 60
tgagttaggt tgggcaaggg ccatcctctc taaacctcaa tttcctcatc tgaactctga 120
gctgcttgac atactgagtt gagattaagg gcaggtgaag caacctttag gtaccaaagt 180
cattcccacc atgcagtcac cttgtcatta cttacacttt tcttcttttt cattttacag 240
taaaaaagtc aagaacatgt aaaaactgtg gcttttctgg aatggaattg gacatagccc 300
aagaacagaa agaaccttgc tggggttgga ggtttcactt gcacatcatg gagggtttag 360
tgcttatcta atttgtgcct cactggactt gtccaattaa tgaagttgat tcatattgca 420
tcatagtttg ctttgtttaa gcatcacatt aaagttaaac tgnattttat gttatttata 480
gctgnaggtt ttctgngttt agctatttaa ta 512

<210> 147
<211> 119
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 21, 36, 72, 76
<223> n = A,T,C or G

<400> 147
ctcaaaatac aatgnttgat nttatgcact ttgtcnctat taacatcctt tttttcatgt 60
agatttcaat anttgngtaa ttttagaagc attattttag gaatatatag ttgtcacag 119

<210> 148
<211> 346
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> 11, 18, 28, 133, 162, 232, 257, 293, 305
<223> n = A,T,C or G

<400> 148
ctagttctgt ncccccanga gacctggntg tgtgtgtgtg agtggttgac cttcctccat 60
cccctggtcc ttcccttccc ttcccgaggc acagagagac agggcaggat ccacgtgccc 120
attgtggagg canagaaaag agaaagtgtt ttatatacgg tncttattta atatcccttt 180
ttaattagaa attaaaacag ttaatttaat taaagagtag ggtttttttt cngtattctt 240
ggttaatatt taatttnaac tatttatgag atgtatcttt tgctctctct tgntctctta 300
tttgnaccgg tttttgtata taaaattcat gtttccaatc tctctc 346

<210> 149
<211> 544
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 411, 505, 513, 515, 533, 539
<223> n = A,T,C or G

<400> 149
ctagttctgt ccccccagga gacctggttg tgtgtgtgtg agtggttgac cttcctccat 60
cccctggtcc ttcccttccc ttcccgaggc acagagagac agggcaggat ccacgtgccc 120
attgtggagg cagagaaaag agaaagtgtt ttatatacgg tacttattta atatcccttt 180
ttaattagaa attaaaacag ttaatttaat taaagagtag ggtttttttt cagtattctt 240
ggttaatatt taatttcaac tatttatgag atgtatcttt tgctctctct tgctctctta 300
tttgtaccgg tttttgtata taaaattcat gtttccaatc tctctctccc tgatcggtga 360
cagtcactag cttatcttga acagatattt aattttgcta acactcagct ntgccctccc 420
cgatcccctg gctccccagc acacattcct ttgaaataag ttttcaatat acatctacat 480
actatatata tatttggcaa cttgnatttg ggngnatata tatatatata tgnttatgna 540
tata 544

<210> 150
<211> 314
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 10, 242, 262
<223> n = A,T,C or G

<400> 150
ctagtccagn gtggtggaat tcaatccttt ttctttttt tggaggtccc accgagatag 60
ataggaactt ggattgctga attcaaaaac agagcccatt cttaagatca cttggtgcct 120
taaagacacg cattccaaag tggaatgtgg ttgaagaaag tgggccaggt ggttgaagaa 180
agccatgtgg gagctcagca aatcccaagg gcttattatg acactccaga tggtctcctt 240
ancatctcag ctcttctgca angaagagct tgggtgttag gcctcagagg ctgtagggtc 300
cttgggttac agag 314

<210> 151
<211> 188
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> 10, 33, 44, 61, 84, 122, 138, 151, 161, 167
<223> n = A,T,C or G

<400> 151
ctagtccagn gtggtggaat tcgcgcagac canacttcgc tcgnactcgt gcgcctcgct 60
ncgcttttcc tccgcaacca tgtntgacaa acccgatatg gctgagatcg agaaattcga 120
tnagtcgaaa ctgaaganga cagagacgca ngagaaaaat ncactgnctt ccaaagaaac 180
gattgaac 188

<210> 152
<211> 487
<212> DNA
<213> Homo sapiens

<400> 152
ctagtccagt gtggtggaat tcgcactccc aaagaactgg gtactcaaca ctgagcagat 60
ctgttctttg agctaaaaac catgtgctgt accaagagtt tgctcctggc tgctttgatg 120
tcagtgctgc tactccacct ctgcggcgaa tcagaagcag caagcaactt tgactgctgt 180
cttggataca cagaccgtat tcttcatcct aaatttattg tgggcttcac acggcagctg 240
gccaatgaag gctgtgacat caatgctatc atctttcaca caaagaaaaa gttgtctgtg 300
tgcgcaaatc caaaacagac ttgggtgaaa tatattgtgc gtctcctcag taaaaaagtc 360
aagaacatgt aaaaactgtg gcttttctgg aatggaattg gacatagccc aagaacagaa 420
agaaccttgc tggggttgga ggtttcactt gcacatcatg gagggtttag tgcttatcta 480
atttgtg 487

<210> 153
<211> 397
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 14, 15, 16, 38, 59, 70, 72, 76, 81, 87, 89, 98, 99, 156,
158, 165, 205, 217, 229, 237, 242, 253, 266, 300, 301, 311,
327, 332, 393
<223> n = A,T,C or G

<400> 153
ctagtccagt gtgnnngaat tcccgaagcg ggagcggnca aaatgaagtt taatccctnt 60
gtgacttccn ancgangcaa naatcgnana aggcattnna atgcaccttc ccacattcga 120
aggaagatta tgtcttcccc tctttccaaa gagctnanac agaantacaa cgtgcgatcc 180
atgcccatcc gaaaggatga tgaanttcag gttgtangtg gacactatna aggtcancaa 240
antggcaaag tantccaggt ttacangaag aaatatgtta tctacattga acgggtgcan 300
ngggaaaagg ntaatggcac aactgtncac gnaggcattc accccagcaa ggtggttatc 360
actaggctaa aactggacaa agaccgcaaa aanatcc 397

<210> 154
<211> 170
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 10, 112
<223> n = A,T,C or G

<400> 154

ccaaaccccn tctgcttctg cccatcacaa gtgccactac cgccatgggc ctcactatct 60
cctccctctt ctcccgacta tttggcaaga agcagatgcg cattttgatg gntggattgg 120
atgctgctgg caagacaacc attctgtata aactgaagtt aggggagata 170

<210> 155
<211> 212
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 190
<223> n = A,T,C or G

<400> 155
tatgagcaag tgaatatgcg gatagaaggc tgtatcattg gttttgatga gtatatgaac 60
cttgtattag atgatgcaga agagattcat tctaaaacaa agtcaagaaa acaactgggt 120
cggatcatgc taaaaggaga taatattact ctgctacaaa gtgtctccaa ctagaaatga 180
tcaatgaagn gagaaattgt tgagaaggat ac 212

<210> 156
<211> 544
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 508
<223> n = A,T,C or G

<400> 156
ctagtttcca aagcggagac ttccgacttc cttacaggat gaggctgggc attgcctggg 60
acagcctatg taaggccatg tgccccttgc cctaacaact cactgcagtg ctcttcatag 120
acacatcttg cagcattttt cttaaggcta tgcttcagtt tttctttgta agccatcaca 180
agccatagtg gtaggtttgc cctttggtac agaaggtgag ttaaagctgg tggaaaaggc 240
ttattgcatt gcattcagag taacctgtgt gcatactcta gaagagtagg gaaaataatg 300
cttgttacaa ttcgacctaa tatgtgcatt gtaaaataaa tgccatattt caaacaaaac 360
acgtaatttt tttacagtat gttttattac cttttgatat ctgttgttgc aatgttagtg 420
atgttttaaa atgtgatcga aaatataatg cttctaagaa ggaacagtag tggaatgaat 480
gtctaaaaga tctttatgtg tttatggnct gcagaaggat ttttgtgatg aaaggggatt 540
tttt 544

<210> 157
<211> 305
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 34, 51, 126, 202, 246, 249, 267, 275
<223> n = A,T,C or G

<400> 157
ctagttagtg cagcttttca ttgtgttgtg tggntgggct cataactagg ntgagttttt 60
ctcctctgct gaggaaacag taccgaagtt ctttttcttg tggcatttgt attataaaaa 120
cttggngtgg gggaggagca caaaactcca gcccactgaa cctctgccaa ttaagatggt 180
gttgggttag gttacatctg gntactgtcc tgggaaaatc atttttatag agatggcctt 240
ccaagnggnt ttaaaattta ctgaagnttt taggncaatt atgtatgttg actaaattta 300

caaat 305

<210> 158
<211> 213
<212> DNA
<213> Homo sapiens

<400> 158
ctagtgagct ctaggctgta gaaatttaaa aactacaatg tgattaactc gagcctttag 60
ttttcatcca tgtacatgga tcacagtttg ctttgatctt cttcaatatg tgaatttggg 120
ctcacagaat caaagcctat gcttggttta atgcttgcaa tctgagctct tgaacaaata 180
aaattaacta ttgtagtgtg aaaaaaaaaa aaa 213

<210> 159
<211> 125
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 32, 38, 104, 116
<223> n = A,T,C or G

<400> 159
atcgccaaga gatcaaagat aaaatctttt gngaaagngt ataactacaa tcacctaatg 60
cccacaaggt actctgtgga tatccccttg gacaaaactg tcgncaataa ggatgncttc 120
agaga 125

<210> 160
<211> 247
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 226
<223> n = A,T,C or G

<400> 160
ctagttagac tctttagaat actccaagag ttagggcagc agagtggagc gatttagaaa 60
gaacatttta aaacaatcag ttaatttacc atgtaaaatt gctgtaaatg ataatgtgta 120
cagattttct gttcaaatat tcaattgtaa acttcttgtt aagactgtta cgtttctatt 180
gcttttgtat gggatattgc aaaaataaaa aggaaagaac cctcanaaaa aaaaaaaaaa 240
aaagggc 247

<210> 161
<211> 373
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 359, 360
<223> n = A,T,C or G

<400> 161
ctagtataga aaataatacg aaactttaaa aagcattgga gtgtcagtat gttgaatcag 60
tagtttcact ttaactgtaa acaatttctt aggacaccat ttgggctagt ttctgtgtaa 120

gtgtaaatac tacaaaaact tatttatact gttcttatgt catttgttat attcatagat 180
ttatatgatg atatgacatc tggctaaaaa gaaattattg caaaactaac cactatgtac 240
ttttttataa atactgtatg gacaaaaaat ggcatttttt atattaaatt gtttagctct 300
ggcaaaaaaa aaaaatttta agagctggta ctaataaagg attattatga ctgttaaann 360
aaaaaaaaaa agg 373

<210> 162
<211> 407
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 17, 19, 21, 180, 227, 232, 382, 388, 401
<223> n = A,T,C or G

<400> 162
ctagtaggat agaaacncng ngtcccgaga gtaaggagag aagctactat tgattagagc 60
ctaacccagg ttaactgcaa gaagaggcgg gatactttca gctttccatg taactgtatg 120
cataaagcca atgtagtcca gtttctaaga tcatgttcca agctaactga atcccacttn 180
aatacacact catgaactcc tgatggaaca ataacaggcc caagccngtg gnatgatgtg 240
cacacttgct agactcagaa aaaatactac tctcataaat gggtgggagt attttggtga 300
caacctactt tgcttggctg agtgaaggaa tgatattcat atattcattt attccatgga 360
catttagtta gtgcttttta tntaccangc atgatgctga ntgacac 407

<210> 163
<211> 396
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 160, 305, 324
<223> n = A,T,C or G

<400> 163
ctagtgtgtc tgatcagtga cttctacccg ggagctgtga cagtggcctg gaaggcagat 60
ggcagccccg tcaaggcggg agtggagacc accaaaccct ccaaacagag caacaacaag 120
tacgcggcca gcagctacct gagcctgacg cccgagcagn ggaagtccca cagaagctac 180
agctgccagg tcacgcatga agggagcacc gtggagaaga cagtggcccc tacagaatgt 240
tcataggttc ccaactctaa ccccacccac gggagcctgg agctgcagga tcccagggga 300
ggggnctctc tccccatccc aagncatcca gcccttctcc ctgcactcat gaaaccccaa 360
taaatatcct cattgacaac caaaaaaaaa aaaaaa 396

<210> 164
<211> 136
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 72
<223> n = A,T,C or G

<400> 164
ctagtccagt gtggtggaat tcaccaaatg gcggatgacg ccggtgcagc gggggggccc 60
ggggggccctg gnggccctgg gatggggaac cgcggtggct tccgcggagg tttcggcagt 120
ggcatccggg gccggg 136

<210> 165
<211> 167
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 19, 20, 21, 50, 90, 116, 117, 131
<223> n = A,T,C or G

<400> 165
ctagtccagt gtggtggann ncctctgtta tttatggtgt gacccccctgn aggtgccctc 60
ggcccaccgg ggctatttat tgtttaattn atttgttgag gttatttttct ctgagnnagt 120
ctgcctctcc naagccccag gggacagtgg ggaggcaggg gaggggg 167

<210> 166
<211> 282
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 22, 23, 25, 81, 82, 194, 236
<223> n = A,T,C or G

<400> 166
ctagtgacaa gctcctggtc tnnanatgtc ttctcgttaa ggagatgggc cttttggagg 60
taaaggataa aatgaatgag nntctgtcat gattcactat tctagaactt gcatgacctt 120
tactgtgtta gctctttgaa tgttcttgaa attttagact ttctttgtaa acaaatgata 180
tgtccttatc atgngtataa aagctgttat gtgcaacagt gtggagattc cttgtntgat 240
ttaataaaat acttaaacac tgaaaaaaaa aaaaaaaagg gc 282

<210> 167
<211> 409
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 377
<223> n = A,T,C or G

<400> 167
ctagtgagcc aggcacatct ggccttggga aactcatcct acaggggaag gccagttttt 60
ttcccttcaa ttcctcaagt ctgggtggtg acaaggtagg ggctaggtac tggactacca 120
caggttttta ggaactaagg tgtttctcat aaacacaaaa tgttgggtga aactgggaac 180
aactactcag aagctaattt atttgcttaa atggaaagtg tgggagccac taccctctct 240
tttgatctgc caaggatttc ctctcagagc tgttgcacag acagagattg tacttggtaa 300
gataccaaac aagacagata tggatctaaa tttctaatgt gttctatggg tttcaattct 360
gaaaaaagga aaatgantaa agattttaat aaataaaaaa aaaaaaaaa 409

<210> 168
<211> 370
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> 359, 360
<223> n = A,T,C or G

<400> 168
ctagtataga aaataatacg aaactttaaa aagcattgga gtgtcagtat gttgaatcag 60
tagtttcact ttaactgtaa acaatttctt aggacaccat ttgggctagt ttctgtgtaa 120
gtgtaaatac tacaaaaact tatttatact gttcttatgt catttgttat attcatagat 180
ttatatgatg atatgacatc tggctaaaaa gaaattattg caaaactaac cactatgtac 240
ttttttataa atactgtatg gacaaaaaat ggcatttttt atattaaatt gtttagctct 300
ggcaaaaaaa aaaaatttta agagctggta ctaataaagg attattatga ctgttaaann 360
aaaaaaaaaa 370

<210> 169
<211> 379
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 359, 360, 373, 378
<223> n = A,T,C or G

<400> 169
ctagtataga aaataatacg aaactttaaa aagcattgga gtgtcagtat gttgaatcag 60
tagtttcact ttaactgtaa acaatttctt aggacaccat ttgggctagt ttctgtgtaa 120
gtgtaaatac tacaaaaact tatttatact gttcttatgt catttgttat attcatagat 180
ttatatgatg atatgacatc tggctaaaaa gaaattattg caaaactaac cactatgtac 240
ttttttataa atactgtatg gacaaaaaat ggcatttttt atattaaatt gtttagctct 300
ggcaaaaaaa aaaaatttta agagctggta ctaataaagg attattatga ctgttaaann 360
aaaaaaaaaa aanaaggnc 379

<210> 170
<211> 222
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 147, 197
<223> n = A,T,C or G

<400> 170
ctagtgagct ctaggctgta gaaatttaaa aactacaatg tgattaactc gagcctttag 60
ttttcatcca tgtacatgga tcacagtttg ctttgatctt cttcaatatg tgaatttggg 120
ctcacagaat caaagcctat gcttggntta atgcttgcaa tctgagctct tgaacaaata 180
aaattaacta ttgtagngtg gaaaaaaaaa aaaaaaaagg gg 222

<210> 171
<211> 298
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 122, 167, 262
<223> n = A,T,C or G

<400> 171
ctagtataga aaataatacg aaactttaaa aagcattgga gtgtcagtat gttgaatcag 60
tagtttcact ttaactgtaa acaatttctt aggacaccat ttgggctagt ttctgtgtaa 120
gngtaaatac tacaaaaact tatttatact gttcttatgt catttgntat attcatagat 180
ttatatgatg atatgacatc tggctaaaaa gaaattattg caaaactaac cactatgtac 240
ttttttataa atactgtatg gncaaaaaat ggcatttttt atattaaatt gtttagct 298

<210> 172
<211> 373
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 20, 22
<223> n = A,T,C or G

<400> 172
ctagtataga aaataatacn anactttaaa aagcattgga gtgtcagtat gttgaatcag 60
tagtttcact ttaactgtaa acaatttctt aggacaccat ttgggctagt ttctgtgtaa 120
gtgtaaatac tacaaaaact tatttatact gttcttatgt catttgttat attcatagat 180
ttatatgatg atatgacatc tggctaaaaa gaaattattg caaaactaac cactatgtac 240
ttttttataa atactgtatg gacaaaaaat ggcatttttt atattaaatt gtttagctct 300
ggcaaaaaaa aaaaatttta agagctggta ctaataaagg attattatga ctgttaaatt 360
aaaaaaaaaa agg 373

<210> 173
<211> 398
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 50, 94, 164, 166, 184, 214, 225, 249, 253, 280, 288,
292, 306, 323
<223> n = A,T,C or G

<400> 173
ctagtccagt gtggnggaat tcgcagcctg aggtgatctg tgaaaatggn tcgctattca 60
cttgacccgg agaaccccac gaaatcatgc aaancaagag gttccaatct tcgtgttcac 120
tttaagaaca ctcgtgaaac tgctcaggcc atcaagggta tgcntntacg aaaagccacg 180
aagnatctga aagatgtcac tttacagaaa cagngtgtac cattncgacg ttacaatggt 240
ggagttggna ggngtgcgca ggccaagcaa tggggctggn cacaaggncg gnggcccaaa 300
aagagngctg aatttttgct gcncatgctt aaaaacgcag agagtaatgc tgaacttaag 360
ggtttagatg tagattctct ggtcattgag catatcca 398

<210> 174
<211> 422
<212> DNA
<213> Homo sapiens

<400> 174
ctagtccagt gtggtggaat tcgcgagaat gaagactatt ctcagcaatc agactgtcga 60
cattccagaa aatgtcgaca ttactctgaa gggacgcaca gttatcgtga agggccccag 120
aggaaccctg cggagggact tcaatcacat caatgtagaa ctcagccttc ttggaaagaa 180
aaaaaagagg ctccgggttg acaaatggtg gggtaacaga aaggaactgg ctaccgttcg 240
gactatttgt agtcatgtac agaacatgat caagggtgtt acactgggct tccgttacaa 300
gatgaggtct gtgtatgctc acttccccat caacgttgtt atccaggaga atgggtctct 360

tgttgaaatc cgaaatttct tgggtgaaaa atatatccgc agggttcgga tgagaccagg 420
tg 422

<210> 175
<211> 470
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 438
<223> n = A,T,C or G

<400> 175
ctagtccatg ggctgagacc ggggcatctc ttttcttcat actgcaatgt tgctagatac 60
atgatcagac accagagggt tgggcattct tgcaatacct taacagtgct gaaatctgca 120
gcatggtact aaggaagtta aagtttgaat gtaaccactt tatttaaaag gttttttct 180
ttaatttaaa tgaaatgggg ttgaagtgaa catgattttg ttgaccatgt tcgtgaatta 240
cagatgcaac atgcattggt agaatcgtgt gatggtcttt tgtgatactt aatttttaca 300
tatcccagtc tctgtatgta tctgcataga caaagaaaaa acaaactcct gctttgcttt 360
tattgaaggg tttccaggac tgcgtgtctg ctcctgagct ctgttttaag gtatgtgtat 420
cctttgcttg tattttgnat taaaaaaaat aagaaaaaag aagcctttat 470

<210> 176
<211> 265
<212> DNA
<213> Homo sapiens

<400> 176
ctagttcttt gtagcagagt acataactac ataatgccaa ctctggaatc aaatttcctt 60
gtttgaatcc tgggacccta ttgcattaaa gtacaaatac tatgtatttt taatctatga 120
tggtttatgt gaataggatt ttctcagttg tcagccatga cttatgttta ttactaaata 180
aacttcaaac tcctgttgaa cattgtgtat aacttagaat aatgaaatat aaggagtatg 240
tgtagaaaaa aaaaaaaaaa agggc 265

<210> 177
<211> 431
<212> DNA
<213> Homo sapiens

<400> 177
ctagtaggat agaaacactg tgtcccgaga gtaaggagag aagctactat tgattagagc 60
ctaacccagg ttaactgcaa gaagaggcgg gatactttca gctttccatg taactgtatg 120
cataaagcca atgtagtcca gtttctaaga tcatgttcca agctaactga atcccacttc 180
aatacacact catgaactcc tgatggaaca ataacaggcc caagcctgtg gtatgatgtg 240
cacacttgct agactcagaa aaaatactac tctcataaat gggtgggagt attttggtga 300
caacctactt tgcttggctg agtgaaggaa tgatattcat atattcattt attccatgga 360
catttagtta gtgctttta tataccaggc atgatgctga gtgacactct tgtgtatatt 420
ttccaaattt t 431

<210> 178
<211> 484
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 350

<223> n = A,T,C or G

<400> 178
ctagtcctct tagaatttct tgcgctttga ttttttttagg gcttgtgccc tgtttcactt 60
atagggtcta gaatgcttgt gttgagtaaa aaggagatgc ccaatattca aagctgctaa 120
atgttctctt tgccataaag actccgtgta actgtgtgaa cacttgggat ttttctcctc 180
tgtcccgagg tcgtcgtctg ctttcttttt tgggtttctt tctagaagat tgagaagtgc 240
atatgacagg ctgagagcac ctccccaaac acacaagctc tcagccacag gcagcttctc 300
cacagcccca gcttcgcaca ggctcctgga gggctgcctg ggggaggcan acatgggagt 360
gccaaggtgg ccagatggtt ccaggactac aatgtcttta tttttaactg tttgccactg 420
ctgccctcac ccctgcccgg ctctggagta ccgtctgccc cagacaagtg ggagtgaaat 480
gggg 484

<210> 179
<211> 592
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 499
<223> n = A,T,C or G

<400> 179
ctagtccagt gtggtggaat tcctaaatca aaggaacttg tttcttcaag ctcttctggc 60
agtgattctg acagtgaggt tgacaaaaag ttaaagagga aaaagcaagt tgctccagaa 120
aaacctgtaa agaaacaaaa gacaggtgag acttcgagag ccctgtcatc ttctaaacag 180
agcagcagca gcagagatga taacatgttt cagattggga aaatgaggta cgttagtgtt 240
cgcgatttta aaggcaaagt gctaattgat attagagaat attggatgga tcctgaaggt 300
gaaatgaaac caggaagaaa aggtatttct ttaaatccag aacaatggag ccagctgaag 360
gaacagattt ctgacattga tgatgcagta agaaaactgt aaaattcgag ccatataaat 420
aaaacctgta ctgttctagt tgttttaatc tgtcttttta cattggcttt tgttttctaa 480
atgttctcca agctattgna tgtttggatt gcagaagaat ttgtaagatg aatacttttt 540
tttaatgtgc attattaaaa atattgagtg aagctaattg tcaactttat ta 592

<210> 180
<211> 199
<212> DNA
<213> Homo sapiens

<400> 180
ctagtccagt gtggtggaat tcgaaggact catgaccaca gtccatgcca tcactgccac 60
ccagaagact gtggatggcc cctccgggaa actgtggcgt gatggccgcg gggctctcca 120
gaacatcatc cctgcctcta ctggcgctgc caaggctgtg ggcaaggtca tccctgagct 180
gaacgggaag ctcactggc 199

<210> 181
<211> 104
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 10, 15, 17, 25, 31, 34, 41, 45, 49, 58, 71
<223> n = A,T,C or G

<400> 181
ctagtccagn gtggngnaat tcctnttgcg ncgncagccg ngccncatng ctcagacncc 60

atggggaagg ngaagggcgg agtcaacgga tttgggcgta ttgg 104

<210> 182
<211> 402
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 175, 193, 196, 197, 206, 236, 299, 377, 382
<223> n = A,T,C or G

<400> 182
ctagtaagca tgacctgggg aaatggtcag accttgtatt gtgtttttgg ccttgaaagt 60
agcaagtgac cagaatctgc catggcaaca ggctttaaaa aagaccctta aaaagacact 120
gtctcaactg tggtgttagc accagccagc tctctgtaca tttgctagct tgtantttтc 180
taagactgag tanacnntct tatttntaga aagtggaggt ctggtttgta actttncttg 240
tacttaattg ggtaaaagtc ttttccacaa accaccatct attttgtgaa ctttgttant 300
catcttttat ttggtaaatt atgaactggt gtaaatttgt acagttcatg tatattgatt 360
gtggcaaagt tgtacangat tnctatattt tggatgagaa at 402

<210> 183
<211> 332
<212> DNA
<213> Homo sapiens

<400> 183
ctagtttgat cgtgatggcg aaacattaga gaaatgcaaa gacatgacca tcataattgt 60
caggagaagg cattggttag gattgggaag cggcaagcag aagcatttag ggattggctg 120
gcaatgtttt acttctcggc tgagtgaggg ttgcatcggt gtttatttga taacacgttc 180
taggggctgg gcaagatggc tcatgtttgt agtctcagta ctttgggagg ccaaagatgg 240
gaggattgct tgagcccgtg agtttgagac cagcgtgggt gacatggcga gaccctgtct 300
ctacaaaaaa ttaaaaaaaa aaaaaaaagg gc 332

<210> 184
<211> 343
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 18, 209, 231, 233, 234, 298, 334, 340
<223> n = A,T,C or G

<400> 184
ctagttagtg cagcttcntc attgtgttgt gtggttggtc tcataactag gttgagtttt 60
tctcctctgc tgaggaaaca gtaccgaagt tctttttctt gtggcatttg tattataaaa 120
acttggtgtg ggggaggagc acaaaactcc agcccactga acctctgcca attaagatgg 180
tgttgggtta ggttacatct ggttactgnc ctgggaaaat catttttata ncnnatggcc 240
ttccaagtgg ttttaaaatt tactgaagtt tttaggtcaa ttatgtatgt tgactaantt 300
tacaaataaa cttgtttatc caaaaaaaaa aaanaaaaan ggc 343

<210> 185
<211> 341
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> 325
<223> n = A,T,C or G

<400> 185
ctagttagtg cagcttttca ttgtgttgtg tggttggtct cataactagg ttgagttttt 60
ctcctctgct gaggaaacag taccgaagtt ctttttcttg tggcatttgt attataaaaa 120
cttggtgtgg gggaggagca caaaactcca gcccactgaa cctctgccaa ttaagatggt 180
gttgggttag gttacatctg gttactgtcc tgggaaaatc atttttatag agatggcctt 240
ccaagtggtt ttaaaattta ctgaagtttt taggtcaatt atgtatgttg actaaattta 300
caaataaact tgtttatcca aaaanaaaaa aaaaaaaggg c 341

<210> 186
<211> 342
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 16, 17, 18, 281
<223> n = A,T,C or G

<400> 186
ctagttagtg cagctnnntc attgtgttgt gtggttggtc tcataactag gttgagtttt 60
tctcctctgc tgaggaaaca gtaccgaagt tctttttctt gtggcatttg tattataaaa 120
acttggtgtg ggggaggagc acaaaactcc agcccactga acctctgcca attaagatgg 180
tgttgggtta ggttacatct ggttactgtc ctgggaaaat catttttata gagatggcct 240
tccaagtggt tttaaaattt actgaagttt ttaggtcaat natgtatgtt gactaaattt 300
acaaataaac ttgtttatcc aaaaaaaaaa aaaaaaaagg gc 342

<210> 187
<211> 132
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 3, 34, 39, 41, 47, 50, 69, 70, 102, 104
<223> n = A,T,C or G

<400> 187
ctngtccagt gtggtggaat tcgcagcctg aggngatcng ngaaaanggn tcgctattca 60
cttgacccnn agaaccccac gaaatcatgc aaatcaagag gntncaatct tcgtgttcac 120
tttaagaaca ct 132

<210> 188
<211> 199
<212> DNA
<213> Homo sapiens

<400> 188
ctagtcacag ccctatactc cctctacata tttaccacaa cacaatgagg ctcactcacc 60
caccacatta acaacataaa accctcattc acacgagaaa acaccctcat gttcatacac 120
ctatccccca ttctcctcct atccctcaac cccgacatca ttaccgggtt ttcctcttaa 180
aaaaaaaaaa aaaaagggc 199

<210> 189
<211> 481

<212> DNA
<213> Homo sapiens

<400> 189
ctagtaggat agaaacactg tgtcccgaga gtaaggagag aagctactat tgattagagc 60
ctaacccagg ttaactgcaa gaagaggcgg gatactttca gctttccatg taactgtatg 120
cataaagcca atgtagtcca gtttctaaga tcatgttcca agctaactga atcccacttc 180
aatacacact catgaactcc tgatggaaca ataacaggcc caagcctgtg gtatgatgtg 240
cacacttgct agactcagaa aaaatactac tctcataaat gggtgggagt attttggtga 300
caacctactt tgcttggctg agtgaaggaa tgatattcat atattcattt attccatgga 360
catttagtta gtgcttttta tataccaggc atgatgctga gtgacactct tgtgtatatt 420
tccaaatttt tgtacagtcg ctgcacatat ttgaaatcat atattaagac tttccaaaga 480
t 481

<210> 190
<211> 351
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 86, 324, 326
<223> n = A,T,C or G

<400> 190
ctagttagtg cagcttttca ttgtgttgtg tggttggtct cataactagg ttgagttttt 60
ctcctctgct gaggaaacag taccgnagtt ctttttcttg tggcatttgt attataaaaa 120
cttggtgtgg gggaggagca caaaactcca gcccactgaa cctctgccaa ttaagatggt 180
gttgggttag gttacatctg gttactgtcc tgggaaaatc atttttatag agatggcctt 240
ccaagtggtt ttaaaattta ctgaagtttt taggtcaatt atgtatgttg actaaattta 300
caaataaact tgtttatcca aaananaaaa aaaaaaaaaa aaaaaagggg c 351

<210> 191
<211> 539
<212> DNA
<213> Homo sapiens

<400> 191
ctagtcacta ctgtcttctc cttgtagcta atcaatcaat attcttccct tgcctgtggg 60
cagtggagag tgctgctggg tgtacgctgc acctgcccac tgagttgggg aaagaggata 120
atcagtgagc actgttctgc tcagagctcc tgatctaccc caccccctag gatccaggac 180
tgggtcaaag ctgcatgaaa ccaggccctg gcagcaacct gggaatggct ggaggtggga 240
gagaacctga cttctctttc cctctccctc ctccaacatt actggaactc tatcctgtta 300
ggatcttctg agcttgtttc cctgctgggt gggacagagg acaaaggaga agggagggtc 360
tagaagaggc agcccttctt tgtcctctgg ggtaaatgag cttgacctag agtaaatgga 420
gagaccaaaa gcctctgatt tttaatttcc ataaaatgtt agaagtatat atatacatat 480
atatatttct ttaaattttt gagtctttga tatgtctaaa aatccattcc ctctgccct 539

<210> 192
<211> 344
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 3, 38, 267, 275, 322
<223> n = A,T,C or G

<400> 192
ctngttagtg cagcttttca ttgtgttgtg tggttggnct cataactagg ttgagttttt 60
ctcctctgct gaggaaacag taccgaagtt ctttttcttg tggcatttgt attataaaaa 120
cttggtgtgg gggaggagca caaaactcca gcccactgaa cctctgccaa ttaagatggt 180
gttgggttag gttacatctg gttactgtcc tgggaaaatc atttttatag agatggcctt 240
ccaagtggtt ttaaaattta ctgaagnttt taggncaatt atgtatgttg actaaattta 300
caaataaact tgtttatcca anaaaaaaaa aaaaaaaaag gggg 344

<210> 193
<211> 469
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 448, 449
<223> n = A,T,C or G

<400> 193
ctagtttgcc agaatattcc aagacatgtt ttagaagcta cctatggcat taacatcata 60
acgcctagag aggatgaaga tccccaccga cctccaacat cggaagaact gttgacagct 120
tatggataca tgcgaggatt catgacagcg catggacagc cagaccagcc tcgatctgcg 180
cgctacatcc tgaaggacta tgtcagtggt aagctgctgt actgccatcc tcctcctgga 240
agagatcctg taacttttca gcatcaacac cagcgactcc tagagaacaa aatgaacagt 300
gatgaaataa aaatgcagct aggcagaaat aaaaaagcaa agcagattga aaatatcgtt 360
gacaaaactt ttttccatca agagaatgtg agggctttga ccaaaggagt ccaggctgtg 420
atgggttaca agcccgggag tggtgtannt gactgcatcc actgcgagc 469

<210> 194
<211> 451
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 247, 249, 262, 386, 393
<223> n = A,T,C or G

<400> 194
ctagtccagt gtggtggaat tcctcaagta caagcctgtc tgcaaccagg tggaatgtca 60
tccttacttc aaccagagaa aactgctgga tttctgcaag tcaaaagaca ttgttctggt 120
tgcctatagt gctctgggat cccatcgaga agaaccatgg gtggacccga actccccggt 180
gctcttggag gacccagtcc tttgtgcctt ggcaaaaaag cacaagcgaa ccccagccct 240
gattgcncnc tgcgctacca gntgcagcgt ggggttgtgg tcctggccaa gagctacaat 300
gagcagcgca tcagacagaa cgtgcaggtg tttgaattcc agttgacttc agaggagatg 360
aaagccatag atggcctaaa cagaanatgt gcnatatttg acccttgata ttttttgctg 420
gccccctaa ttatccattt tctgatgaat a 451

<210> 195
<211> 322
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 36, 173, 189, 287
<223> n = A,T,C or G

<400> 195
ctagtccagt gtggtggaat tcggaaactg tggcgngatg gccgcggggc tctccagaac 60
atcatccctg cctctactgg cgctgccaag gctgtgggca aggtcatccc tgagctgaac 120
gggaagctca ctggcatggc cttccgtgtc cccactgcca acgtgtcagt ggnggacctg 180
acctgccgnc tagaaaaacc tgccaaatat gatgacatca agaaggtggt gaagcaggcg 240
tcggagggcc ccctcaaggg catcctgggc tacactgagc accaggnggg ctcctctgac 300
ttcaacagcg acacccactc ct 322

<210> 196
<211> 490
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 470
<223> n = A,T,C or G

<400> 196
ctagtccagt gtggtggaat tccgcctcgg aggcgttcag ctgcttcaag atgaagctga 60
acatctcctt cccagccact ggctgccaga aactcattga agtggacgat gaacgcaaac 120
ttcgtacttt ctatgagaag cgtatggcca cagaagttgc tgctgacgct ctgggtgaag 180
aatggaaggg ttatgtggtc cgaatcagtg gtgggaacga caaacaaggt ttccccatga 240
agcagggtgt cttgacccat ggccgtgtcc gcctgctact gagtaagggg cattcctgtt 300
acagaccaag gagaactgga gaaagaaaga gaaaatcagt tcgtggttgc attgtggatg 360
caaatctgag cgttctcaac ttggttattg taaaaaaagg agagaaggat attcctggac 420
tgactgatac tacagtgcct cgccgcctgg gccccaaaag gagctagcan aatccgcaaa 480
cttttcaatc 490

<210> 197
<211> 327
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 76, 136, 177, 191, 226, 248, 307, 311
<223> n = A,T,C or G

<400> 197
ctagtgcttt acctttatta atgaactgtg acaggaagcc caaggcagtg ttcctcacca 60
ataacttcag agaagncagt tggagaaaat gaagaaaaag gctggctgaa aatcactata 120
accatcagtt actggnttca gttgacaaaa tatataatgg gttactgctg tcattgncca 180
tgcctacaga naatttattt tgtatttttg aataaaaaac atttgnacat tcctgatact 240
gggtacanga gccatgtacc agtgtactgc tttcaactta aatcactgag gcatttttac 300
tactatnctg ntaaaatcag gatttta 327

<210> 198
<211> 202
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 9, 22, 39, 45, 61, 66, 67, 119, 120, 179, 194
<223> n = A,T,C or G

<400> 198

gtttcacang gatcctctga ancctctct gtgccccang tacanatgcc attacttctg 60
ntttcnnatc tcctcaggca aaagtggagg gtgccttatg ggccctcctc ataggttgnn 120
tctgcataca cgaacctaac ccaaatttgc tttggtgcca gaaaaactga gctatgttng 180
aacaaagatg tcgngcaaac tg 202

<210> 199
<211> 485
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 391
<223> n = A,T,C or G

<400> 199
ttacctttat taatgaactg tgacaggaag cccaaggcag tgttcctcac caataacttc 60
agagaagtca gttggagaaa atgaagaaaa aggctggctg aaaatcacta taaccatcag 120
ttactggttt cagttgacaa aatatataat ggtttactgc tgtcattgtc catgcctaca 180
gataatttat tttgtatttt tgaataaaaa acatttgtac attcctgata ctgggtacaa 240
gagccatgta ccagtgtact gctttcaact taaatcactg aggcattttt actactattc 300
tgttaaaatc aggattttag tgcttgccac caccagatga gaagttaagc agcctttctg 360
tggagagtga gaataattgt gtacaaagta ngagaagtat ccaattatgt gacaaccttt 420
gtgtaataaa aatttgttta aagttaaaaa aaaaaaaaaa gggcggccgc caccgcggtg 480
gagct 485

<210> 200
<211> 196
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 9, 15, 16, 26, 42, 48, 49, 160
<223> n = A,T,C or G

<400> 200
ccagtgtgnt ggaannccgg cgttgntctg gattcccgtc gnaacttnna gggaaacttt 60
cacaatgtcc ggagcccttg atgtcctgca aatgaaggag gaggatgtcc ttaagttcct 120
tgcagcagga acccacttag gtggcaccaa tcttgacttn cagatggaac agtacatcta 180
taaaaggaaa agtgat 196

<210> 201
<211> 91
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 40
<223> n = A,T,C or G

<400> 201
ttatgaggat atgcatttaa ttttaaattt tataatttan attcagcatg aattgcaata 60
aatggatcat cagcgggttt aaacgggccc t 91

<210> 202
<211> 367

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 175, 220
<223> n = A,T,C or G

<400> 202
tggaattcgc cgagcaggag gcgccatcat gggagtggac atccgccata acaaggaccg 60
aaaggttcgg cgcaaggagc ccaagagcca ggatatctac ctgaggctgt tggtcaagtt 120
atacaggttt ctggccagaa gaaccaactc cacattcaac caggttgtgt tgaanaggtt 180
tgtttatgag tcgcaccaac cggccgcctc tgtccctttn ccggatgatc cggaagatga 240
agcttcctgg ccgggaaaac aagacggccg tggttgtggg gaccataact gatgatgtgc 300
gggttcagga ggtacccaaa ctgaaggtat gtgcactgcg cgtgaccagc cgggcccgca 360
gccgcat 367

<210> 203
<211> 213
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 2
<223> n = A,T,C or G

<400> 203
nngagctcta ggctgtagaa atttaaaaac tacaatgtga ttaactcgag cctttagttt 60
tcatccatgt acatggatca cagtttgctt tgatcttctt caatatgtga atttgggctc 120
acagaatcaa agcctatgct tggtttaatg cttgcaatct gagctcttga acaaataaaa 180
ttaactattg tagtgtgaaa aaaaaaaaaa aaa 213

<210> 204
<211> 94
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1
<223> n = A,T,C or G

<400> 204
naatttcgtg tatatgaatc tttctcgaag atctggtcaa aactgtattc agtttcctgc 60
ccagaatgat cagattgaag gtggttggtt ttta 94

<210> 205
<211> 520
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 10, 11, 92, 272, 331, 342, 354, 420, 429, 449, 462, 475,
492, 493, 498
<223> n = A,T,C or G

<400> 205
tggaattccn nagactgagc ggttgtggcc gcgttgccga cctccagcag cagtcggctt 60
ctctacgcag aacccgggag taggagactc anaatcgaat ctcttctccc tccccttctt 120
gtgagatttt tttgatcttc agctacattt tcggctttgt gagaaacctt accatcaaac 180
acgatggcca gcaacgttac caacaagaca gatcctcgct ccatgaactc ccgtgtattt 240
cattgggaac ctcaacactc ttgtggttca anaaatctga tgtggaggca atcttttcga 300
agtatggcaa aattgtgggc tgctctgttc ntaagggctt tnccttcgtt cagnatgtta 360
atgagagaaa tgcccgggct gctgtagcag gagaggatgg caggaatgat tgctggccan 420
gtttttagnt attaacctgg ctgcagagnc caaaagtgaa cngaggaaaa agcangtgtg 480
aaacgatctg tnncgganat gtacggctcc tcttttgact 520

<210> 206
<211> 84
<212> DNA
<213> Homo sapiens

<400> 206
ccttaagaag tcatgattaa cttatgaaaa aattatttgg ggacaggagt gtgatacctt 60
ccttggtttt tttttgcagc cctc 84

<210> 207
<211> 125
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 35, 74, 87, 88, 100, 101
<223> n = A,T,C or G

<400> 207
tcgagcggcc gccctttttt ttttttttnt tttgntttga ggatatgcat ttaattttaa 60
attttataat ttanattcag catgaanngc aataaatggn ncatcagcgg gtttaaacgg 120
gccct 125

<210> 208
<211> 212
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 2
<223> n = A,T,C or G

<400> 208
nngagctcta ggctgtagaa atttaaaaac tacaatgtga ttaactcgag cctttagttt 60
tcatccatgt acatggatca cagtttgctt tgatcttctt caatatgtga atttgggctc 120
acagaatcaa agcctatgct tggtttaatg cttgcaatct gagctcttga acaaataaaa 180
ttaactattg tagtgtgaaa aaaaaaaaaa aa 212

<210> 209
<211> 270
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<222> 189, 190
<223> n = A,T,C or G

<400> 209
gacaagctcc tggtcttgag atgtcttctc gttaaggaga tgggcctttt ggaggtaaag 60
gataaaatga atgagttctg tcatgattca ctattctaga acttgcatga cctttactgt 120
gttagctctt tgaatgttct tgaaatttta gactttcttt gtaaacaaat gatatgtcct 180
tatcattgnn taaaagctgt tatgtgcaac agtgtggaga ttccttgtct gatttaataa 240
aatacttaaa cactgaaaaa aaaaaaaaaa 270

<210> 210
<211> 415
<212> DNA
<213> Homo sapiens

<400> 210
aggccttcca gttcactgac aaacatgggg aagtgtgccc agctggctgg aaacctggca 60
gtgataccat caagcctgat gtccaaaaga gcaaagaata tttctccaag cagaagtgag 120
cgctgggctg ttttagtgcc aggctgcggt gggcagccat gagaacaaaa cctcttctgt 180
atttttttt tccattagta aaacacaaga cttcagattc agccgaattg tggtgtctta 240
caaggcaggc ctttcctaca gggggtggag agaccagcct ttcttccttt ggtaggaatg 300
gcctgagttg gcgttgtggg caggctactg gtttgtatga tgtattagta gagcaaccca 360
ttaatctttt gtagtttgta ttaaacttga actgagaaaa aaaaaaaaaa aaaaa 415

<210> 211
<211> 234
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 54, 55, 163, 176, 192, 215, 218, 230
<223> n = A,T,C or G

<400> 211
actgaaaaga gccatgctgt ctagtcttga agtccctcat ttaaacagag gtcnngcaat 60
aggcgcctgg cagtgtcaag cctgaaacca agcaataccg tcatgtttca gccaagccca 120
gagccctaag attacaaaca actatggccg gaacctcctc agntctccct ctgcanagtt 180
ccctacccta anagaatgtt accacctgaa cagtnctngg tgaatctgan agga 234

<210> 212
<211> 531
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 2, 3, 460
<223> n = A,T,C or G

<400> 212
nnncaaaaat gctaaaataa tttgggagaa aatatttttt aagtagtgtt atagtttcat 60
gtttatcttt tattatgttt tgtgaagttg tgtcttttca ctaattacct atactatgcc 120
aatatttcct tatatctatc cataacattt atactacatt tgtaagagaa tatgcacgtg 180
aaacttaaca ctttataagg taaaaatgag gtttccaaga tttaataatc tgatcaagtt 240
cttgttattt ccaaatagaa tggacttggt ctgttaaggg ctaaggagaa gaggaagata 300
aggttaaaag ttgttaatga ccaaacattc taaaagaaat gcaaaaaaaa agtttatttt 360
caagccttcg aactatttaa ggaaagcaaa atcatttcct aaatgcatat catttgtgag 420

aatttctcat taatatcctg aatcattcat ttcagctaan gcttcatgtt gactcgatat 480
gtcatctagg aaagtactat ttcatggtcc aaacctgttg ccatagttgg t 531

<210> 213
<211> 229
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 28, 61, 62
<223> n = A,T,C or G

<400> 213
gataagcttg atatcgaatt cctgcagncc gggggatcca ctagtaggat agaaacactg 60
nntcccgaga gtaaggagag aagctactat tgattagagc ctaacccagg ttaactgcaa 120
gaagaggcgg gatactttca gctttccatg taactgtatg cataaagcca atgtagtcca 180
gtttctaaga tcatgttcca agctaactga atcccacttc aatacacac 229

<210> 214
<211> 196
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 2, 73, 79
<223> n = A,T,C or G

<400> 214
nnttaccttt attaatgaac tgtgacagga agcccaaggc agtgttcctc accaataact 60
tcagagaagt canttggana aaatgaagaa aaaggctggc tgaaaatcac tataaccatc 120
agttactggt ttcagttgac aaaatatata atggtttact gctgtcattg tccatgccta 180
cagataattt attttg 196

<210> 215
<211> 213
<212> DNA
<213> Homo sapiens

<400> 215
aattcctgca gcccggggga tccactagtc cagtgtggtg gaattccccg agcgccgctc 60
cggctgcacc gcgctcgctc cgagtttcag gctcgtgcta agctagcgcc gtcgtcgtct 120
cccttcagtc gccatcatga ttatctaccg ggacctcatc agccacgatg agatgttctc 180
cgacatctac aagatccggg agatcgcgga cgg 213

<210> 216
<211> 161
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 14, 15, 17, 103
<223> n = A,T,C or G

<400> 216
tttggcttaa attnngnctt ttgaagttga atgcttaatc ccgggaaaga ggaacaggag 60

tgccatactc ctggtctttc cagtttagaa aaggctctgt gcncaaggag ggaccacagg 120
agctgggacc tgcctgcccc tgtcttttcc ccttggtttt g 161

<210> 217
<211> 417
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 48, 49, 384, 392
<223> n = A,T,C or G

<400> 217
ttacctttat taatgaactg tgacaggaag cccaaggcag tgttcctnnc caataacttc 60
agagaagtca gttggagaaa atgaagaaaa aggctggctg aaaatcacta taaccatcag 120
ttactggttt cagttgacaa aatatataat ggtttactgc tgtcattgtc catgcctaca 180
gataatttat tttgtatttt tgaataaaaa acatttgtac attcctgata ctgggtacaa 240
gagccatgta ccagtgtact gctttcaact taaatcactg aggcattttt actactattc 300
tgttaaaatc aggattttag tgcttgccac caccagatga gaagttaagc agcctttctg 360
tggagagtga gaataattgt tgtncaaagt anagaagtat ccaattatgt gacaacc 417

<210> 218
<211> 425
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 18, 19, 31, 250, 251, 290
<223> n = A,T,C or G

<400> 218
cagtgtggtg gaattcgnng ttgaaaactg naattgaaca ggtttacgca aatggcatcc 60
ggaacattga ccttcactat attgtgttac tgcggaaatg caaaacttag tccatcggcg 120
gatttatcca tttttactga tggtcgtggt attgatggca attttgtcct tccaagtccg 180
ccagtttaag cgcctttatg aacatattaa aaatgacaag taccttgtgg gtcaacgact 240
cgtgaactan naacggaaat ctggcaaaca aggctcatct ccaccacctn cacagtcatc 300
ccaagaataa agtagtttgt ctcaacaact tgaccttccc ctttacatgt ccttttttgt 360
ggacttctct ctttggagat ttttcccagt gatctctcag ccgttgtttt taagttaaat 420
gtatt 425

<210> 219
<211> 470
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 422
<223> n = A,T,C or G

<400> 219
aattccatcg atggcatttc agtctatagg taaacttcct ggaagctgga tttggagaca 60
gtttatcatc tgattattgg gctttcgtat aggtccttag ggagcagctt acctgaaatg 120
catttagtgt acaccagtct gtaaacttca acctgtaatg aaagtgtaat aaatgtacat 180
tgagttgatg tgataatgtg atataataag aaatatatat ttgatcttcc tatctagttc 240
cttgttcaga gctcctaaaa cccttgtaat ttccaaagtg atggagtaca tcttttgttc 300

tagtatttgg tctttgaccc cagttcctga cacaaagctc ctaaattcct ttaaatttcc 360
cagtgatagg agaatttttt gttctaatga ggtcactctt gatgggcacc tggataactc 420
angatggggg ctgctcacaa agaccacatc atgattggaa gtttcaaact 470

<210> 220
<211> 536
<212> DNA
<213> Homo sapiens

<400> 220
aaaaagcagc attgccaaat aatccctaat tttccactaa aaatataatg aaatgatgtt 60
aagctttttg aaaagtttag gttaaaccta ctgttgttag attaatgtat ttgttgcttc 120
cctttatctg gaatgtggca ttagcttttt tattttaacc ctctttaatt cttattcaat 180
tccatgactt aaggttggag agctaaacac tgggattttt ggataacaga ctgacagttt 240
tgcataatta taatcggcat tgtacataga aaggatatgg ctacctttg ttaaatctgc 300
actttctaaa tatcaaaaaa gggaaatgaa gtataaatca atttttgtat aatctgtttg 360
aaacatgagt tttatttgct taatattagg gctttgcccc ttttctgtaa gtctcttggg 420
atcctgtgta gaagctgttc tcattaaaca ccaaacagtt aagtccattc tctggtacta 480
gctacaaatt cggtttcata ttctacttaa caatttaaat aaactgaaat atttct 536

<210> 221
<211> 384
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 5, 6, 355, 359
<223> n = A,T,C or G

<400> 221
ntccnntgtg gtggaattcc ttttcaattt gaatcccata tggggagaca gaggacgaaa 60
cagccatcct gtcgacttct ttgtaagggg catcagagtc aaagactgcc agaacaccca 120
cactgatcct acctgcataa tgtggaatga atgctatgga taaactgctg aagatggttc 180
ctgtccattt gactctgaag ggtgtcttct ttcacgttga agaacaggag acaatcaaaa 240
tgtgaaacgt atgctgaagc caaccagaac atcaaaggac agtcaaaagc gctaaccatg 300
aaactatatt tctactaata cattctttta aaaaaaaaat aaaaacaaac ctgcntgtnc 360
gtgaaaaaaa aaaaaaaaag ggcg 384

<210> 222
<211> 212
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 11
<223> n = A,T,C or G

<400> 222
tggaattcgc ngttgaaaac tgtaattgaa caggtttacg caaatggcat ccggaacatt 60
gaccttcact atattgtgtt actgcggaaa tgcaaaactt agtccatcgg cggatttatc 120
catttttact gatggtcgtg gtattgatgg caattttgtc cttccaagtc cgccagttta 180
agcgccttta tgaacatatt aaaaatgaca ag 212

<210> 223
<211> 304
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<222> 141
<223> n = A,T,C or G

<400> 223
ctgctgatag aaagcactat acatcctatt gtttctttct ttccaaaatc agccttctgt 60
ctgtaacaaa aatgtacttt atagagatgg aggaaaaggt ctaatactac atagccttaa 120
gtgtttctgt cattgttcaa ntgtattttc tgtaacagaa acatatttgg aatgtttttc 180
ttttcccctt ataaattgta attcctgaaa tactgctgct ttaaaaagtc ccactgtcag 240
attaataatt atctaacaat tgaatattgt aaatatactt gtcttacctc tcaataaaag 300
ggta 304

<210> 224
<211> 101
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 4, 15
<223> n = A,T,C or G

<400> 224
gtcnccgaga gtgangagag aagctactat tgattagagc ctaacccagg ttaactgcaa 60
gaagaggcgg gatactttca gctttccatg taactgtatg c 101

<210> 225
<211> 442
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 407, 418, 433
<223> n = A,T,C or G

<400> 225
ctagtccagt gtggtggaat tctgagtcct tgatttcaaa gttttgttgt acttaaatgg 60
taataagcac tgtaaacttc tgcaacaagc atgcagcttt gcaaacccat taaggggaag 120
aatgaaagct gttccttggt cctagtaaga agacaaactg cttcccttac tttgctgagg 180
gtttgaataa acctaggact tccgagctat gtcagtacta ttcaggtaac actagggcct 240
tggaaattcc tgtactgtgt ctcatggatt tggcactagc caaagcgagg cacccttact 300
ggcttacctc ctcatggcag cctactctcc ttgagtgtat gagtagccag ggtaaggggt 360
aaaaggatag taagcataga aaccactaga aagtgggctt aatgganttc ttgtggcnct 420
cagctcaatg canttagctg aa 442

<210> 226
<211> 437
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 347, 349
<223> n = A,T,C or G

<400> 226
ctagtccagt gtggtggaat tcacgacctg tctcgccgag cgcacgcctt gccgccgccc 60
cgcagaaatg cttcggttac ccacagtctt tcgccagatg agaccggtgt ccagggtact 120
ggctcctcat ctcactcggg cttatgccaa agatgtaaaa tttggtgcag atgcccgagc 180
cttaatgctt caaggtgtag acctttttagc cgatgctgtg gccgttacaa tggggccaaa 240
gggaagaaca gtgattattg agcagagttg gggaagtccc aaagtaacaa aagatggtgt 300
gactgttgca aagtcaattg acttaaaaga taaatacaag aacattngna gctaaacttg 360
ttcaagatgt tgccaataac acaaatgaag aagctgggga tggcactacc actgctactg 420
tactggcacg ctctata 437

<210> 227
<211> 382
<212> DNA
<213> Homo sapiens

<400> 227
ctagtttaag gagactggcc gaacctctgc ccaaacaatc tgtggatgga aaagcaccac 60
ttgctactgg agaggatgat gatgatgaag ttccagatct tgtggagaat tttgatgagg 120
cttccaagaa tgaggcaaac tgaattgagt caacttctga agataaaacc tgaagaagtt 180
actgggagct gctattttat attatgactg ctttttaaga aatttttgtt tatggatctg 240
ataaaatcta gatctctaat atttttaagc ccaagcccct tggacactgc agctcttttc 300
agtttttgct tatacacaat tcattctttg cagctaatta agccgaagaa gcctgggaat 360
caagtttgaa acaaagatta at 382

<210> 228
<211> 346
<212> DNA
<213> Homo sapiens

<400> 228
ctagtggaag attaccggcg tgttattgaa cgacttgctc aagagtaaag attatactgc 60
tctgtacagg aagcttgcaa attttctgta caatgtgctg tgaaaaatct gatgacttta 120
attttaaaat cttgtgacat tttgcttata ctaaaagtta tctatcttta gttgaatatt 180
ttcttttgga gagattgtat attttaaaat actgtttaga gtttatgagc atatattgca 240
tttaaagaaa gataaagctt ctgaaatact actgcaattg cttcccttct taaacagtat 300
aataaatgct tagttgtgat atgttaaaaa aaaaaaaaaa aagggc 346

<210> 229
<211> 340
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 265, 269, 336
<223> n = A,T,C or G

<400> 229
ctagttattt actttcctcc gcttcagaaa gtttttcaga ctgagagcct aagcatactg 60
gatctgttgt ttcttttggg tctcacctca tcagtgtgca tagtggcaga aattataaag 120
aaggttgaaa ggagcaggga aaagatccag aagcatgtta gttcgacatc atcatctttt 180
cttgaagtat gatgcatatt gcattatttt atttgcaaac taggaattgc agtctgagga 240
tcatttagaa gggcaagttc aagangatnt gaagatttga gaacttttta actattcatt 300
gactaaaaat gaacattaat gttaaagact taaganttta 340

<210> 230
<211> 348

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 188, 264, 265, 324
<223> n = A,T,C or G

<400> 230
ctagtccagt gtggtggaat tcgcatcatg gaggtttgaa gatgccgcat ttggattgga 60
tgaattccaa attctgcttg cttgcttttt aatattgata tgcttataca cttacacttt 120
atgcacaaaa tgtagggtta taataatgtt aacatggaca tgatcttctt tataattcta 180
ctttgagngc tgtctccatg tttgatgtat ctgagcaggt tgctccacag gtagctctag 240
gagggctggc gacttagagg tggnnagcag agaattctct tatccaacat caacatcttg 300
gtcagatttg aactcttcaa tctnttgcac tcaaagcttg ttaagata 348

<210> 231
<211> 360
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 224, 264, 286, 314
<223> n = A,T,C or G

<400> 231
ctagtaagca tgacctgggg aaatggtcag accttgtatt gtgtttttgg ccttgaaagt 60
agcaagtgac cagaatctgc catggcaaca ggctttaaaa aagaccctta aaaagacact 120
gtctcaactg tggtgttagc accagccagc tctctgtaca tttgctagct tgtagttttc 180
taagactgag taaacttctt atttttagaa agtggaggtc tggnttgtaa ctttccttgt 240
acttaattgg gtaaaagtct tttncacaaa ccaccatcta ttttgngaac tttgttagtc 300
atcttttatt tggnaaatta tgaactggtg taaatttgta cagttcatgt atattgattg 360

<210> 232
<211> 214
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 16, 34, 67, 74, 87, 138, 145, 146, 149, 183, 187
<223> n = A,T,C or G

<400> 232
ctctgtgctc cgcggngacc cagacgaggc tcgngacttt gcagccggcc ttagtgctcg 60
cgcaggntcc tggnagagtt acacagntgt gccgccagta tagcgacatg cctcctttga 120
cgttagaggg catccagnac cgtgnnctnt acgtattgaa actctatgac aagattgacc 180
canagangct ttcagtaaat tctcatttta tgaa 214

<210> 233
<211> 457
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<222> 171, 386
<223> n = A,T,C or G

<400> 233
ctagtgtaac tccttcatgc aataaactga aaagagccat gctgtctagt cttgaagtcc 60
ctcatttaaa cagaggtcaa gcaataggcg cctggcagtg tcaagcctga aaccaagcaa 120
taccgtcatg tttcagccaa gcccagagcc ctaagattac aaacaactat ngccggaacc 180
tcctcagctc tccctctgca gagttcccta ccctaagaga atgttaccac ctgaacagtc 240
ctcggtgaat ctgagaggag aggatggggt aaggcagaag caccagctgt actactagaa 300
gggagctttt ggtggtagat cccctggtgt ctccaacctg actaggtgga cagagctcaa 360
agaggccctc ttaccgctag cgaggngata ggacatctgg cttgccacaa aggtctgttc 420
gaccagacat atcctagcta agggatgtcc aaacatc 457

<210> 234
<211> 342
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 34, 89, 148, 267
<223> n = A,T,C or G

<400> 234
ctagttagtg cagcttttca ttgtgttgtg tggntggtct cataactagg ttgagttttt 60
ctcctctgct gaggaaacag taccgaagnt ctttttcttg tggcatttgt attataaaaa 120
cttggtgtgg gggaggagca caaaactnca gcccactgaa cctctgccaa ttaagatggt 180
gttgggttag gttacatctg gttactgtcc tgggaaaatc atttttatag agatggcctt 240
ccaagtggtt ttaaaattta ctgaagnttt taggtcaatt atgtatgttg actaaattta 300
caaataaact tgtttatcca aaaaaaaaaa aaaaaaaaaa gg 342

<210> 235
<211> 332
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 38, 274
<223> n = A,T,C or G

<400> 235
ctagttagtg cagcttttca ttgtgttgtg tggttggnct cataactagg ttgagttttt 60
ctcctctgct gaggaaacag taccgaagtt ctttttcttg tggcatttgt attataaaaa 120
cttggtgtgg gggaggagca caaaactcca gcccactgaa cctctgccaa ttaagatggt 180
gttgggttag gttacatctg gttactgtcc tgggaaaatc atttttatag agatggcctt 240
ccaagtggtt ttaaaattta ctgaagtttt tagntcaatt atgtatgttg actaaattta 300
caaataaact tgtttatcca aaaaaaaaaa aa 332

<210> 236
<211> 323
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 276
<223> n = A,T,C or G

<400> 236
ctagtccagt gtggtggaat tcgtctcatt ctgacttcat ggagaattaa tcccaccttt 60
aagcaaaggc tactaagtta atggtatttt ctgtgcagaa attaaatttt attttcagca 120
tttagcccag gaattcttcc agtaggtgct cagctattta aaaacaaaac tattctcaaa 180
cattcatcat tagacaactg gagtttttgc tggttttgta acctaccaaa atggataggc 240
tgtttgaaca ttccacattc aaaagttttg tagggnggtg ggaaatgggg gatcttcaat 300
gtttatttta aaataaaata aaa 323

<210> 237
<211> 377
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 264, 286
<223> n = A,T,C or G

<400> 237
ctagtaagca tgacctgggg aaatggtcag accttgtatt gtgtttttgg ccttgaaagt 60
agcaagtgac cagaatctgc catggcaaca ggctttaaaa aagaccctta aaaagacact 120
gtctcaactg tggtgttagc accagccagc tctctgtaca tttgctagct tgtagttttc 180
taagactgag taaacttctt atttttagaa agtggaggtc tggtttgtaa ctttccttgt 240
acttaattgg gtaaaagtct tttncacaaa ccaccatcta ttttgngaac tttgttagtc 300
atcttttatt tggtaaatta tgaactggtg taaatttgta cagttcatgt atattgattg 360
tggcaaagtt gtacaga 377

<210> 238
<211> 105
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 103
<223> n = A,T,C or G

<400> 238
ctagttgatg tatggtatct ttagatattt gcctgtctgt ttgctcaaaa ttgcttctaa 60
aacaataaag attcttttat ttcttaaaaa aaaaaaaaaa aangg 105

<210> 239
<211> 218
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 16
<223> n = A,T,C or G

<400> 239
gagctctagg ctgtanaaat ttaaaaacta caatgtgatt aactcgagcc tttagttttc 60
atccatgtac atggatcaca gtttgctttg atcttcttca atatgtgaat ttgggctcac 120
agaatcaaag cctatgcttg gtttaatgct tgcaatctga gctcttgaac aaataaaatt 180
aactattgta gtgtgaaaac aaaaaaaaaa aaaaaggg 218

<210> 240
<211> 279
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 179, 263
<223> n = A,T,C or G

<400> 240
ctagtgacaa gctcctggtc ttgagatgtc ttctcgttaa ggagatgggc cttttggagg 60
taaaggataa aatgaatgag ttctgtcatg attcactatt ctagaacttg catgaccttt 120
actgtgttag ctctttgaat gttcttgaaa ttttagactt tctttgtaaa caaatgatnt 180
gtccttatca ttgtataaaa gctgttatgt gcaacagtgt ggagattcct tgtctgattt 240
aataaaatac ttaaacactg aanaaaaaaa aaaaagggc 279

<210> 241
<211> 271
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 19, 30, 56, 61, 67, 151, 168, 183, 195, 249, 255
<223> n = A,T,C or G

<400> 241
ctagtgacaa gctcctggnc ttgagatgtn ttctcgttaa ggagatgggc cttttngagg 60
naaaggntaa aatgaatgag ttctgtcatg attcactatt ctagaacttg catgaccttt 120
actgtgttag ctctttgaat gttcttgaaa ntttagactt tctttgtnaa caaatgatat 180
gtncttatca ttgtntaaaa gctgttatgt gcaacagtgt ggagattcct tgtctgattt 240
aataaaatnc ttaancactg aaaaaaaaaa a 271

<210> 242
<211> 345
<212> DNA
<213> Homo sapiens

<400> 242
ctagtccagt gtggtggaat tcgcctcgga ggcgttcagc ttgcttcaag atgaagctga 60
acatctcctt cccagccact ggctgccaga aactcattga agtggacgat gaacgcaaac 120
ttcgtacttt ctatgagaag cgtatggcca cagaagttgc tgctgacgct ctgggtgaag 180
aatggaaggg ttatgtggtc cgaatcagtg gtgggaacga caaacaaggt ttccccatga 240
agcaagggtg tcttgaccca tggccgtgtc cgcctgctac tgagtaaggg gcattcctgt 300
tacagaccaa ggagaactgg agaaagaaag agaaaatcag ttcgt 345

<210> 243
<211> 418
<212> DNA
<213> Homo sapiens

<400> 243
ctagtttaag gagactggcc gaagctctgc ccaaacaatc tgtggatgga aaagcaccac 60
ttgctactgg agaggatgat gatgatgaag ttccagatct tgtggagaat tttgatgagg 120
cttccaagaa tgaggcaaac tgaattgagt caacttctga agataaaacc tgaagaagtt 180
actgggagct gctattttat attatgactg ctttttaaga aatttttgtt tatggatctg 240
ataaaatcta gatctctaat atttttaagc ccaagcccct tggacactgc agctcttttc 300

agtttttgct tatacacaat tcattctttg cagctaatta agccgaagaa gcctgggaat 360
caagtttgaa acaaagatta ataaagttct ttgcctagta aaaaaaaaaa aaaagggc 418

<210> 244
<211> 350
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 177, 213, 278
<223> n = A,T,C or G

<400> 244
ctagtccagt gtggtggaat tcgtctcatt ctgacttcat ggagaattaa tcccaccttt 60
aagcaaaggc tactaagtta atggtatttt ctgtgcagaa attaaatttt attttcagca 120
tttagcccag gaattcttcc agtaggtgct cagctattta aaaacaaaac tattctnaaa 180
cattcatcat tagacaactg gagtttttgc tgnttttgta acctaccaaa atggataggc 240
tgttgaacat tccacattca aaagttttgt agggtggngg gaaatggggg atcttcaatg 300
tttattttaa aataaaataa aataagttct tgacttttaa aaaaaaaaaa 350

<210> 245
<211> 419
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 394, 401
<223> n = A,T,C or G

<400> 245
ctagtaaaaa gcagcattgc caaataatcc ctaattttcc actaaaaata taatgaaatg 60
atgttaagct ttttgaaaag tttaggttaa acctactgtt gttagattaa tgtatttgtt 120
gcttcccttt atctggaatg tggcattagc ttttttattt taaccctctt taattcttat 180
tcaattccat gacttaaggt tggagagcta aacactggga tttttggata acagactgac 240
agttttgcat aattataatc ggcattgtac atagaaagga tatggctacc ttttgttaaa 300
tctgcacttt ctaaatatca aaaaagggaa atgaagtata aatcaatttt tgtataatct 360
gtttgaaaca tgagttttat ttgcttaata ttanggcttt ncccctttc tgtaagtct 419

<210> 246
<211> 434
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 234, 353, 362, 419
<223> n = A,T,C or G

<400> 246
ctagtaaaaa gcagcattgc caaataatcc ctaattttcc actaaaaata taatgaaatg 60
atgttaagct ttttgaaaag tttaggttaa acctactgtt gttagattaa tgtatttgtt 120
gcttcccttt atctggaatg tggcattagc ttttttattt taaccctctt taattcttat 180
tcaattccat gacttaaggt tggagagcta aacactggga tttttggata acanactgac 240
agttttgcat aattataatc ggcattgtac atagaaagga tatggctacc ttttgttaaa 300
tctgcacttt ctaaatatca aaaaagggaa atgaagtata aatcaatttt tgnataatct 360
gnttgaaaca tgagttttat tttgcttaat attagggctt tgccccttt ctgtaagtnt 420

cttgggatcc tgtg 434

<210> 247
<211> 221
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 218
<223> n = A,T,C or G

<400> 247
ctagtgagct ctaggctgta gaaatttaaa aactacaatg tgattaactc gagcctttag 60
ttttcatcca tgtacatgga tcacagtttg ctttgatctt cttcaatatg tgaatttggg 120
ctcacagaat caaagcctat gcttggttta atgcttgcaa tctgagctct tgaacaaata 180
aaattaacta ttgtagtgtg aaaaaaaaaa aaaaaaangg g 221

<210> 248
<211> 217
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 201
<223> n = A,T,C or G

<400> 248
ctagtgagct ctaggctgta gaaatttaaa aactacaatg tgattaactc gagcctttag 60
ttttcatcca tgtacatgga tcacagtttg ctttgatctt cttcaatatg tgaatttggg 120
ctcacagaat caaagcctat gcttggttta atgcttgcaa tctgagctct tgaacaaata 180
aaattaacta ttgtagtgtg naaaaaaaaa aaaaaaa 217

<210> 249
<211> 357
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 30, 43, 76, 92, 93, 143, 166, 195, 205, 233, 291, 324
<223> n = A,T,C or G

<400> 249
ctagtaggat agaaacactg tgtcccgagn gtaaggagag aanctactat tgattagagc 60
ctaacccagg ttaacnagca agaagaggcg gnntactttc agctttccat gtaactgtat 120
gcataaagcc aatgtagtcc agnttctaag atcatgttcc aagctnactg aatcccactt 180
caatacacac tcatnaactc ctganggaac aataacaggc ccaagcctgt ggnatgatgt 240
gcacacttgc tagactcaga aaaaatacta ctctcataaa tgggtgggag nattttggtg 300
acaacctact ttgcttggct gagngaagga atgatattca tatattcatt tattcca 357

<210> 250
<211> 219
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> 14
<223> n = A,T,C or G

<400> 250
ctagtgagct ctangctgta gaaatttaaa aactacaatg tgattaactc gagcctttag 60
ttttcatcca tgtacatgga tcacagtttg ctttgatctt cttcaatatg tgaatttggg 120
ctcacagaat caaagcctat gcttggttta atgcttgcaa tctgagctct tgaacaaata 180
aaattaacta ttgtagtgtg aaaaaaaaaa aaaaagggc 219

<210> 251
<211> 199
<212> DNA
<213> Homo sapiens

<400> 251
ctagtccagt gtggtggaat tcggccaagg tgcaacttcc ttcggtcgtc ccgaatccgg 60
gttcatccga caccagccgc ctccaccatg ccgccgaagt tcgaccccaa cgagatcaaa 120
gtcgtatacc tgaggtgcac cggaggtgaa gtcggtgcca cttctgccct ggcccccaag 180
atcggccccc tgggtctgt 199

<210> 252
<211> 221
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 218
<223> n = A,T,C or G

<400> 252
ctagtgagct ctaggctgta gaaatttaaa aactacaatg tgattaactc gagcctttag 60
ttttcatcca tgtacatgga tcacagtttg ctttgatctt cttcaatatg tgaatttggg 120
ctcacagaat caaagcctat gcttggttta atgcttgcaa tctgagctct tgaacaaata 180
aaattaacta ttgtagtgtg aaaaaaaaaa aaaaaaangg g 221

<210> 253
<211> 457
<212> DNA
<213> Homo sapiens

<400> 253
ctagtccagt gtggtggaat tcataacatt ccaatcacta ttgtatatat gtgcatgtat 60
tttttaaatt aaagatgtct agttgctttt tataagacca agaaggagaa aatccgacaa 120
cctggaaaga tttttgtttt cactgcttgt atgatgtttc ccattcatac acctataaat 180
ctctaacaag aggccctttg aactgccttg tgttctgtga gaaacaaata tttacttaga 240
gtggaaggac tgattgagaa tgttccaatc caaatgaatg catcacaact tacaatgctg 300
ctcattgttg tgagtactat gagattcaaa tttttctaac atatggaaag ccttttgtcc 360
tccaaagatg agtactaggg atcatgtgtt taaaaaaaga aaggctacga tgactgggca 420
agaagaaaga tgggaaactg aataaagcag ttgatca 457

<210> 254
<211> 391
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> 351, 362, 372, 378
<223> n = A,T,C or G

<400> 254
ctagtgttct ttcagtaaag tacaaagtgt ttattttaca aaagagtagg tactcttgag 60
agcaattcaa atcatgctga caaggatact gatagaaaaa gtgatttctt cttattataa 120
agtacattta aagttcaagg actaacctta tttatttggg aaaggggagg aggaaggaaa 180
tgatatggta cccagacact gggctaggct gcaactttat ctcatttaat actcccagct 240
gtcatgtgag aaagaaagca ggctaggcat gtgaaatcac tttcatggat tattaatgga 300
tttaagaggg catcaatcag ctcaactcaa gatttcataa tcatttttag natttagatt 360
gngcctcaaa gntgtagnac ctcacaatac c 391

<210> 255
<211> 556
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 521, 539
<223> n = A,T,C or G

<400> 255
ctagtcccaa cgcgtttgca aatattcccc tggtagccta cttccttacc cccgaatatt 60
ggtaagatcg agcaatggct tcaggacatg ggttctcttc tcctgtgatc attcaagtgc 120
tcactgcatg aagactggct tgtctcagtg tttcaacctc accagggctg tctcttggtc 180
cacacctcgc tccctgttag tgccgtatga cagcccccat caaatgacct tggccaagtc 240
acggtttctc tgtggtcaag gttggttggc tgattggtgg aaagtagggt ggaccaaagg 300
aggccacgtg agcagtcagc accagttctg caccagcagc gcctccgtcc tagtgggtgt 360
tcctgtttct cctggccctg ggtgggctag ggcctgattc gggaagatgc ctttgcaggg 420
aggggaggat aagtgggatc taccaattga ttctggcaaa acaatttcta agatttttt 480
gctttatgtg ggaaacagat ctaaatctca ttttatgctg nattttatat cttagttgng 540
tttgaaaacg ttttga 556

<210> 256
<211> 212
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 5, 15, 147
<223> n = A,T,C or G

<400> 256
ctagngagct ctagnctgta gaaatttaaa aactacaatg tgattaactc gagcctttag 60
ttttcatcca tgtacatgga tcacagtttg ctttgatctt cttcaatatg tgaatttggg 120
ctcacagaat caaagcctat gcttggntta atgcttgcaa tctgagctct tgaacaaata 180
aaattaacta ttgtagtgtg gaaaaaaaaa aa 212

<210> 257
<211> 459
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<222> 439
<223> n = A,T,C or G

<400> 257
ctagtagtca gttgggagtg gttgctatac cttgacttca tttatatgaa tttccacttt 60
attaaataat agaaaagaaa atcccggtgc ttgcagtaga gtgataggac attctatgct 120
tacagaaaat atagccatga ttgaaatcaa atagtaaagg ctgttctggc tttttatctt 180
cttagctcat cttaaataag cagtacactt ggatgcagtg cgtctgaagt gctaatcagt 240
tgtaacaata gcacaaatcg aacttaggat ttgtttcttc tcttctgtgt ttcgattttt 300
gatcaattct ttaattttgg aagcctataa tacagtttc tattcttgga gataaaaatt 360
aaatggatca ctgatatttt agtcattctg cttctcatct aaatatttcc atattctgta 420
ttaggagaaa attacccctnc cagcaccagc ccccctctc 459

<210> 258
<211> 406
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 368, 405
<223> n = A,T,C or G

<400> 258
ctagtccagt gtggtggaat tccatggagg gtgtagaaga gaagaagaag gaggttcctg 60
ctgtgccaga aacccttaag aaaaagcgaa ggaatttcgc agagctgaag atcaagcgcc 120
tgagaaagaa gtttgcccaa aagatgcttc gaaaggcaag gaggaagctt atctatgaaa 180
aagcaaagca ctatcacaag gaatataggc agatgtacag aactgaaatt cgaatggcga 240
ggatggcaag aaaagctggc aacttctatg tacctgcaga acccaaattg gcgtttgtca 300
tcagaatcag aggtatcaat ggagtgagcc caaaggttcg aaaggtgttg cagcttcttc 360
gccttcgnca aatctccaat ggaacctttg tgaagctcaa caagnc 406

<210> 259
<211> 394
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 385
<223> n = A,T,C or G

<400> 259
ctagtccagt gtggtggaat tcgtcctgcg cggttgttct ctggagcagc gttcttttat 60
ctccgtccgc cttctctcct acctaagtgc gtgccgccac ccgatggaag attcgatgga 120
catggacatg agccccctga ggccccagaa ctatcttttc ggttgtgaac taaaggccga 180
caaagattat cactttaagg tggataatga tgaaaatgag caccagttat ctttaagaac 240
ggtcagttta ggggctggtg caaaggatga gttgcacatt gttgaagcag aggcaatgaa 300
ttacgaaggc agtccaatta aagtaacact ggcaactttg aaaatgtctg tacagccaac 360
ggtttccctt gggggctttg aaatnacacc acca 394

<210> 260
<211> 364
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<222> 295
<223> n = A,T,C or G

<400> 260
ctagtataga aaataatacg aaactttaaa aagcattgga gtgtcagtat gttgaatcag 60
tagtttcact ttaactgtaa acaatttctt aggacaccat ttgggctagt ttctgtgtaa 120
gtgtaaatac tacaaaaact tatttatact gttcttatgt catttgttat attcatagat 180
ttatatgatg atatgacatc tggctaaaaa gaaattattg caaaactaac cactatgtac 240
ttttttataa atactgtatg gacaaaaaat ggcatttttt atattaaatt gtttnagctc 300
tggcaaaaaa aaaaaatttt aagagctggt actaataaag gattattatg actgttaaaa 360
aaaa 364

<210> 261
<211> 458
<212> DNA
<213> Homo sapiens

<400> 261
ctagtagcag gtagagcatg aatgacagca tattatacca tcaagatgtt cttagagcag 60
tgtatggatg gatcgattgt actgccatca gttgtgactg acgttgtatt caaggagaaa 120
gagaaacttg tttagaaagc actttgaaag ttttttgagt acgggggtgc cctgtatcac 180
cccgttatgg ttgaactttc tccttcaaaa ttaccagact tggcagcagt ggcaaattat 240
tgggctaaaa gacttaatca gacatattct gggttcaagg ctcctaatat aatacctggt 300
gcaaacatta tacttccact cattcagatg gttgcatcct gccaggcatc cagtgggact 360
gggaatatgg acacttgaac attaaacatc ctgaagaatt ttggaatgac aggttacaag 420
tgaacataat cagttctcta tattaaaaaa aaaaaaaa 458

<210> 262
<211> 282
<212> DNA
<213> Homo sapiens

<400> 262
ctagtgacaa gctcctggtc ttgagatgtc ttctcgttaa ggagatgggc cttttggagg 60
taaaggataa aatgaatgag ttctgtcatg attcactatt ctagaacttg catgaccttt 120
actgtgttag ctctttgaat gttcttgaaa ttttagactt tctttgtaaa caaatgatat 180
gtccttatca ttgtataaaa gctgttatgt gcaacagtgt ggagattcct tgtctgattt 240
aataaaatac ttaaacactg aaaaaaaaaa aaaaaaaagg gc 282

<210> 263
<211> 278
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 276
<223> n = A,T,C or G

<400> 263
ctagtgacaa gctcctggtc ttgagatgtc ttctcgttaa ggagatgggc cttttggagg 60
taaaggataa aatgaatgag ttctgtcatg attcactatt ctagaacttg catgaccttt 120
actgtgttag ctctttgaat gttcttgaaa ttttagactt tctttgtaaa caaatgatat 180
gtccttatca ttgtataaaa gctgttatgt gcaacagtgt ggagattcct tgtctgattt 240
aataaaatac ttaaacactg aaaaaaaaaa aaaaangg 278

<210> 264
<211> 232

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 28, 209
<223> n = A,T,C or G

<400> 264
ctagtcctac ctctgccact aatgaggngt ttggaggagg taccagccat ataatagggg 60
gtgtatgtgt gaattttgtt taaactctac tgtatattga aatgaaattc atttatttgt 120
cttgacaatg ttcaaatgat gtagattgtc ttagaatgaa tattcataag tactcagaac 180
tcttaagatg cagatgccac ccgtgaggng ctaaattcct aatgtgtatt gt 232

<210> 265
<211> 203
<212> DNA
<213> Homo sapiens

<400> 265
ctagtcacag ccctatactc cctctacata tttaccacaa cacaatgggg ctcactcacc 60
caccacatta acaacataaa accctcattc acacgagaaa acaccctcat gttcatacac 120
ctatccccca ttctcctcct atccctcaac cccgacatca ttaccgggtt ttcctcttaa 180
aaaaaaaaaa aaaaaaaagg ggg 203

<210> 266
<211> 226
<212> DNA
<213> Homo sapiens

<400> 266
ctagtgagct ctaggctgta gaaatttaaa aactacaatg tgattaactc gagcctttag 60
ttttcatcca tgtacatgga tcacagtttg ctttgatctt cttcaatatg tgaatttggg 120
ctcacagaat caaagcctat gcttggatta atgcttgcaa tctgagctct tgaacaaata 180
aaattaacta ttgtagtgtg aaaaaaaaaa aaaaaaaaaa aagggg 226

<210> 267
<211> 325
<212> DNA
<213> Homo sapiens

<400> 267
ctagtttttc ctatcatgtt aacctctgct tttatctcag atgttaaaat aaatggtttg 60
gtgcttttta taaaaagata atctcagtgc tttcctcctt cactgtttca tctaagtgcc 120
tcacattttt ttctacctat aacactctag gatgtatatt ttatataaag tattcttttt 180
ctttttaaa ttaatatctt tctgcacaca aatattattt gtgtttccta aatccaacca 240
ttttcattaa ttcaggcata ttttaactcc actgcttacc tactttcttc aggtaaaggg 300
caaataatga tcgaaaaaaa aaaaa 325

<210> 268
<211> 217
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 79
<223> n = A,T,C or G

<400> 268
ctagtccagt gtggtggaat tctagaagtc tggtttataa aaaagccaaa agtgatggaa 60
tttattccat ttgtcttang aaggcccata atacttgttt ttcttacatg tgactagcaa 120
ctttctccac ttaaagacta aatacctctt tatatgatgt aaattattct aattcatttt 180
aaaatctttt aggtcagcaa aaaaaaaaaa aaagggc 217

<210> 269
<211> 315
<212> DNA
<213> Homo sapiens

<400> 269
ctagtgaaga aaaagaaatt ctgatacggg acaaaaatgc tcttcaaaac atcattcttt 60
atcacctgac accaggagtt ttcattggaa aaggatttga acctggtgtt actaacattt 120
taaagaccac acaaggaagc aaaatctttc tgaaagaagt aaatgataca cttctggtga 180
atgaattgaa atcaaaagaa tctgacatca tgacaacaaa tggtgtaatt catgttgtag 240
ataaactcct ctatccagca gacacacctg ttggaaatga tcaactgctg gaaatactta 300
ataaattaat caaat 315

<210> 270
<211> 412
<212> DNA
<213> Homo sapiens

<400> 270
ctagtgcttc ccagtacttg catggggttc actatttata gttttcttgg gagtatcaca 60
ggaaaatcac aattacacca ctttagaccc tatgtgtagc aggtcacaac ttacccttgt 120
gtgtttagat gtgtatgaaa tacctgtata cgttagtgaa agctgtttac tgtaacgggg 180
aaaaccagat tctttgcatc tgggccctct actgattgtt aaaggagttc ctgtcacctg 240
ctccccccac ccccgcatgc gtctgtccac ttggctaact tttaatatgt gtatttttac 300
attatgtata ttcttaactg gactgtctcg tttagactgt atacatcata tctgacatta 360
ttgtaactac cgtgtgatca gtaagattcc tgtaagaaat actgcttttt aa 412

<210> 271
<211> 218
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 174, 175, 206
<223> n = A,T,C or G

<400> 271
gagctctagg ctgtagaaat ttaaaaacta caatgtgatt aactcgagcc tttagttttc 60
atccatgtac atggatcaca gtttgctttg atcttcttca atatgtgaat ttgggctcac 120
agaatcaaag cctatgcttg gtttaatgct tgcaatctga gctcttgaac aaannaaaat 180
taactattgt agtgtgaaaa aaaaanaaaa aaaagggc 218

<210> 272
<211> 398
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 253

<223> n = A,T,C or G

<400> 272
ctagtccagt gtggtggaat tcgagagcac cgcccagcag ccagtgggtt cccgcgcgtg 60
ccgagactct gaggccttgc acccccacga tcccgtacga tggccgtcaa gaagatcgcg 120
atcttcggcg ccactggcca gaccgggctc accaccctgg cgcaggcggt gcaagcaggt 180
tacgaagtga cagtgctggt gcgggactcc tccaggctgc catcagaggg gccccggccg 240
gcccacgtgg tantgggaga tgttctgcag gcagccgatg tggacaagac cgtggctggg 300
caggacgctg tcatcgtgct gctgggcacc cgcaatgacc tcagtcccac gacagtgatg 360
tccgagggcg cccggaacat tgtggcagcc atgaaggc 398

<210> 273
<211> 496
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 390
<223> n = A,T,C or G

<400> 273
ctagtccagt gtggtggaat tcgcttcctc ctcctcggcc tcaccattcc agaccaaaat 60
tgaaaaaatg gttgacctca cccaggtaat ggatgatgaa gtattcatgg cttttgcatc 120
ctatgcaaca attattcttt caaaaatgat gcttatgagt actgcaactg cattctatag 180
attgacaaga aaggtttttg ccaatccaga agactgtgta gcatttggca aaggagaaaa 240
tgccaagaag tatcttcgaa cagatgacag agtagaacgt gtacgcagag cccacctgaa 300
tgaccttgaa aatattattc catttcttgg aattggcctc ctgtattcct tgagtggtcc 360
cgacccctct acagccatcc tgcacttcan actatttgtc ggagcacgga tctaccacac 420
cattgcatat ttgacacccc ttccccagcc aaatagagct ttgagttttt ttgttggata 480
tggagttact ctttcc 496

<210> 274
<211> 403
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 69, 147
<223> n = A,T,C or G

<400> 274
ctagttaaac atggnctgcg tgccttaaga gagacgcttc ctgcagaaca ggacctgact 60
acaaagaang tttccattgg aattgttggt aaagacttgg agtttacaat ctatgatgat 120
gatgatgtgt ctccattcct ggaaggnctt gaagaaagac cacagagaaa ggcacagcct 180
gctcaacctg ctgatgaacc tgcagaaaag gctgatgaac caatggaaca ttaagtgata 240
agccagtcta tatatgtatt atcaaatatg taagaataca ggcaccacat actgatgaca 300
ataatctata ctttgaacca aaagttgcag agtggtggaa tgctatgttt taggaatcag 360
tccagatgtg agtttttcc aagcaacctc actgaaacct ata 403

<210> 275
<211> 277
<212> DNA
<213> Homo sapiens

<400> 275
ctagtgacaa gctcctggtc ttgagatgtc ttctcgttaa ggagatgggc cttttggagg 60

taaaggataa aatgaatgag ttctgtcatg attcactatt ctagaacttg catgaccttt 120
actgtgttag ctctttgaat gttcttgaaa ttttagactt tctttgtaaa caaatgatat 180
gtccttatca ttgtataaaa gctgttatgt gcaacagtgt ggagattcct tgtctgattt 240
aataaaatac ttaaacactg aaaaaaaaaa aaaaaaa 277

<210> 276
<211> 285
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 65, 228, 230, 247, 249, 264
<223> n = A,T,C or G

<400> 276
ctagtctcag gcttcaacat cgaatacgcc gcaggcccct tcgccctatt cttcatagcc 60
gaatncacaa acattattat aataaacacc ctcaccacta caatcttcct aggaacaaca 120
tatgacgcac tctcccctga actctacaca acatattttg ttcctaggaa gattgtagtg 180
gtgacctccc tgttcttatg aattcgaaca gcataccccc gattccgntn cgaccaactc 240
atacacntnc tatgaaaaaa cttnctacca ctcaccctag catta 285

<210> 277
<211> 188
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 23, 24, 45, 185
<223> n = A,T,C or G

<400> 277
cctatggaaa aaaccaagct tcnntagaat gtctgcctta ctggnttccc cagggaagga 60
aaaatacact tccacccttt tttctaagtg ttcgtcttta gttttgattt tggaaagatg 120
ttaagcattt atttttagtt aaaaataaaa actaatttca tactatttaa aaaaaaaaaa 180
aaaanggg 188

<210> 278
<211> 309
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 19, 71, 72, 129, 181, 190, 203, 210
<223> n = A,T,C or G

<400> 278
ctagttagca tgccagagnc tcgttcgtta tcggaattaa ccagacaaat cgctccacca 60
actaagaacg nncatgcacc accacccacg gaatcgagaa agagctatca atctgtcaat 120
cctgtccgng tccgggccgg gtgaggtttc ccgtgttgag tcaaattaag ccgcaggctc 180
nactcctggn ggtgcccttc cgncaattcn tttaagtttc agctttgcaa ccatactccc 240
cccggaaccc aaagactttg gtttcccgga agctgcccgg cgggtcatgg gaataacgcc 300
gccgcatcg 309

<210> 279
<211> 369

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 142, 154, 155, 217, 338, 364
<223> n = A,T,C or G

<400> 279
ctagtccagt gtggnggaat tccttcgctc gtactcgtgc gcctcgcttc gcttttcctc 60
cgcaaccatg tctgacaaac ccgatatggc tgagatcgag aaattcgata agtcgaaact 120
gaagaagaca gagacgcaag anaaaaatcc actnncttcc aaagaaacga ttgaacagga 180
gaagcaagca ggcgaatcgt aatgaggcgt gcgccgncaa tatgcactgt acattccaca 240
agcattgcct tcttatttta cttcttttag ctgtttaact ttgtaagatg caaagaggtt 300
ggatcaagtt taaatgactg tgctgcccct ttcacatnaa agaactactg acaacgaagg 360
ccgngcctg 369

<210> 280
<211> 509
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 272, 393, 398, 406, 452
<223> n = A,T,C or G

<400> 280
ctagtgaatg aagaacgaac gctggaagta gaaatagagc ctggggtgag agacggcatg 60
gagtacccct ttattggaga aggtgagcct cacgtggatg gggagcctgg agatttacgg 120
ttccgaatca aagttgtcaa gcacccaata tttgaaagga gaggagatga tttgtacaca 180
aatgtgacaa tctcattagt tgagtcactg gttggctttg agatggatat tactcacttg 240
gatggtcaca aggtacatat ttcccgggat angatcacca ggccaggagc gaagctatgg 300
aagaaagggg aagggctccc caactttgac aacaacaata tcaagggctc tttgataatc 360
acttttgatg tggattttcc aaaagaacag ttnacagngg aagcgngaga aggtatcaaa 420
cagctactga aacaagggtc agtgcagaag gnatacaatg gactgcaagg atattgagag 480
tgaataaaat tgggactttg tttaaaaat 509

<210> 281
<211> 526
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 102, 165, 433, 461, 503
<223> n = A,T,C or G

<400> 281
ctagtccagt gtggtggaat tccgccggtg cagcgggggg gcccgggggc cctggtggcc 60
ctgggatggg gaaccgcggt ggcttccgcg gaggtttcgg cngtggcatc cggggccggg 120
gtcgcggccg tggacggggc cggggccgag gccgcggagc tcgcngaggc aaggccgagg 180
ataaggagtg gatgcccgtc accaagttgg gccgcttggt caaggacatg aagatcaagt 240
ccctggagga gatctatctc ttctccctgc ccattaagga atcagagatc attgatttct 300
tcctgggggc ctctctcaag gatgaggttt tgaagattat gccagtgcag aagcagaccc 360
gtgccggcca gcgcaccagg ttcaaggcat ttgttgctat cggggactac aatggccacg 420
tcggtctggg tgnttaagtg ctccaaggag gtggccaccg ncatccgtgg ggccatcatc 480
ctggccaagc tctccatcgt ccncgtgcgc agaggctact ggggga 526

<210> 282
<211> 610
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 342
<223> n = A,T,C or G

<400> 282
ctagtccagt gtggtggaat tcggaagcgc tccgctgtac ctggatcctg ctcctctggg 60
ttgaaacccg ggcgccgcca agatgccggc ttaccactct tctctcatgg atcctgatac 120
caaactcatc ggaaacatgg cactgttgcc tatcagaagt caattcaaag gacctgcccc 180
cagagagaca aaagatacag atattgtgga tgaagccatc tattacttca aggccaatgt 240
cttcttcaaa aactatgaaa ttaagaatga agctgatagg accttgatat atataactct 300
ctacatttct gaatgtctga agaaactgca aaagtgcaat tncaaaagcc aaggtgagaa 360
agaaatgtat acgctgggaa tcactaattt tcccattcct ggagagcctg gttttccact 420
taacgcaatt tatgccaaac ctgcaaacaa acaggaagat gaagtgatga gagcctattt 480
acaacagcta aggcaagaga ctggactgag actttgtgag aaagttttcg accctcagaa 540
tgataaaccc agcaagtggt ggacttgctt tgtgaagaga cagttcatga acaagagtct 600
ttcaggacct 610

<210> 283
<211> 324
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 163, 221, 242
<223> n = A,T,C or G

<400> 283
ctagtctgct gatagaaagc actatacatc ctattgtttc tttctttcca aaatcagcct 60
tctgtctgta acaaaaatgt actttataga gatggaggaa aaggtctaat actacatagc 120
cttaagtgtt tctgtcattg ttcaagtgta ttttctgtaa canaaacata tttggaatgt 180
ttttcttttc cccttataaa ttgtaattcc tgaaatactg ntgctttaaa aagtcccact 240
gncagattat attatctaac aattgaatat tgtaaatata cttgtcttac ctctcaataa 300
aagggtactt ttctattaaa aaaa 324

<210> 284
<211> 437
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 406
<223> n = A,T,C or G

<400> 284
ctagttctgg tacttgtgtc tttgtatgat caaagcatgc aataagcaat acaaaatacc 60
aagccttata cttaaaagaa gtttaacata ttggttaata tactggttaa tatactggtt 120
aaacatattg aatgtatata agtggcaaaa ctagattttt aaggaagtgt acattataat 180
attggagctc agtactgcat gaagagactt cattaaaact aagaaaacat ttatttgggg 240
agaaatttta ggcatttaag aacttgtatt tttctatttt aaaaagttaa attattccgt 300

aatttggaag aagtttcgtt gaatgtagga cataaccgtt tgaagggttt tcatttgaaa 360
aattgatgta ttttgtgcct taatattttg ttcttttaat aaaaangctc tgaatttgaa 420
aaaaaaaaaa aaagggc 437

<210> 285
<211> 503
<212> DNA
<213> Homo sapiens

<400> 285
ctagtccagt gtggtggaat tccagcattc gggccgagat gtctcgctcc gtggccttag 60
ctgtgctcgc gctactctct ctttctggcc tggaggctat ccagcgtact ccaaagattc 120
aggtttactc acgtcatcca gcagagaatg gaaagtcaaa tttcctgaat tgctatgtgt 180
ctgggtttca tccatccgac attgaagttg acttactgaa gaatggagag agaattgaaa 240
aagtggagca ttcagacttg tctttcagca aggactggtc tttctatctc ttgtactaca 300
ctgaattcac ccccactgaa aaagatgagt atgcctgccg tgtgaaccat gtgactttgt 360
cacagcccaa gatagttaag tgggatcgag acatgtaagc agcatcatgg aggtttgaag 420
atgccgcatt tggattggat gaattccaaa ttctgcttgc ttgcttttta atattgatat 480
gcttatacac ttacacttta tgc 503

<210> 286
<211> 374
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 52, 67, 97, 98, 111, 115, 130, 140, 242, 298, 352, 365
<223> n = A,T,C or G

<400> 286
ccgccgcaac ttcaattacc gacgcagacg cccagaaaac cctaaaccac angatggcaa 60
agagacnaaa gcagccgatc caccagctga gaattcnncc gctcccgagg ntgancaggg 120
cggggctgan taaatgccgn cttaccatct ctaccatcat tccggtttag tcatccaaca 180
agaagaaata tgaaattcca gcaataagaa atgaacaaaa gattggagct gaagacctaa 240
antgcttgct ttttgcccgt tgaccagata aatagaacta tctgcattat ctatgcanca 300
tggggttttt attattttta cctaaagacg tctcttttg gtaataacaa angtgttttt 360
taaanaagcc tggt 374

<210> 287
<211> 453
<212> DNA
<213> Homo sapiens

<400> 287
ctagtctgtg tgggactgta cacactttat ttacttcgtt ttggttaagt tggcttctgt 60
ttctagttga ggagtttcct aaaagttcat aacagtgcca ttgtctttat atgaacatag 120
actagagaaa ccgtcctctt tttccatcat aattctaatc taacaatgga agatttgccc 180
atttacactt ttgagacttt ttggtggatg taaataaccc cattctttgc ttgaacacag 240
tattttccca atagcacttt cattgccagt gtctttcttt ggtgcctttc ctgttcagca 300
ttcttagcct gtggcagtaa agagaaactt tgtgctacat gacgacaaag ctgctaaatc 360
tcctattttt ttaaaatcac taacattata ttgcaatgaa ggaaataaaa aagtctctat 420
ttaaattctt ttttaaaaaa aaaaaaaaag ggc 453

<210> 288
<211> 459
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 4, 15, 20, 23, 42, 49, 53, 68, 85, 93, 177, 190, 198, 215,
243, 255, 258, 316, 357, 388, 389
<223> n = A,T,C or G

<400> 288
ctantccagt gtggnggaan tcngacgctc tcagctctcg gngcacggnc cancttcctt 60
caaaatgnct actgttcacg aaatnctgtg cangctcagc ttggagggtg atcactctac 120
acccccaagt gcatatgggt ctgtcaaagc ctatactaac tttgatgctg agcgggntgc 180
tttgaacatn gaaacagnca tcaagaccaa aggtntggat gaggtcacca ttgtcaacat 240
ttngaccaac cgcancantg cacagagaca ggatattgcc ttcgcctacc agagaaggac 300
caaaaaggaa cttgcntcag cactgaagtc agccttatct ggccacctgg agacggngat 360
tttgggccta ttgaagacac ctgctcanna tgacgcttct gagctaaaag cttccatgaa 420
ggggctggga accgacgagg actctctcat tgagatcat 459

<210> 289
<211> 577
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 488
<223> n = A,T,C or G

<400> 289
ctagtgacta attttccctt acagttcctg cttggtccca cccactgaag tagctcatcg 60
tagtgcgggc cgtattagag gcagtggggt acgttagact cagatggaaa agtattctag 120
gtgccagtgt taggatgtca gttttacaaa ataatgaagc aattagctat gtgattgaga 180
gttattgttt ggggatgtgt gttgtggttt tgctttttt ttttagactg tattaataaa 240
catacaacac aagctggcct tgtgttgctg gttcctattc agtatttcct ggggattgtt 300
tgctttttaa gtaaaacact tctgacccat agctcagtat gtctgaattc cagaggtcac 360
atcagcatct ttctgctttg aaaactctca cagctgtggc tgcttcactt agatgcagtg 420
agacacatag ttggtgttcc gattttcaca tccttccatg tatttatctt gaagagataa 480
gcacaganga gaaggtgctc actaacagag gtacattact gcaatgttct cttaacagtt 540
aaacaagctg tttacagttt aaactgctga atattat 577

<210> 290
<211> 404
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 20, 169, 364, 367, 393
<223> n = A,T,C or G

<400> 290
ctagtccagt gtggtggaan tccaaatggc ggatgacgcc ggtgcagcgg gggggcccgg 60
gggccctggt ggccctggga tggggaaccg cggtggcttc cgcggaggtt tcggcagtgg 120
catccggggc cggggtcgcg gccgtggacg gggccggggc cgaggccgng gagctcgcgg 180
aggcaaggcc gaggataagg agtggatgcc cgtcaccaag ttgggccgct tggtcaagga 240
catgaagatc aagtccctgg aggagatcta tctcttctcc ctgcccatta aggaatcaga 300
gatcattgat ttcttcctgg gggcctctct caaggatgag gttttgaaga ttatgccagt 360
gcanaancag acccgtgccg gccagcgcac cangttcaag gcat 404

<210> 291
<211> 383
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 379
<223> n = A,T,C or G

<400> 291
ctagtataga aaataatacg aaactttaaa aagtattgga gtgtcagtat gttgaatcag 60
tagtttcact ttaactgtaa acaatttctt aggacaccat ttgggctagt ttctgtgtaa 120
gtgtaaatac tacaaaaact tatttatact gttcttatgt catttgttat attcatagat 180
ttatatgatg atatgacatc tggctaaaaa gaaattattg caaaactaac cactatgtac 240
ttttttataa atactgtatg gacaaaaaat ggcatttttt atattaaatt gtttagctct 300
ggcaaaaaaa aaaaatttta agagctggta ctaataaagg attattatga ctgttaaaaa 360
aaaaaaaaaa aaaaaaaang ggc 383

<210> 292
<211> 612
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 558, 566, 567
<223> n = A,T,C or G

<400> 292
ctagtgtgct catcctgaac tgttactcca aatccactcc gtttttaaag caaaattatc 60
ttgtgatttt aagaaaagag ttttctattt atttaagaaa gtaacaatgc agtctgcaag 120
ctttcagtag ttttctagtg ctatattcat cctgtaaaac tcttactacg taaccagtaa 180
tcacaaggaa agtgtcccct ttgcatattt ctttaaaatt ctttctttgg aaagtatgat 240
gttgataatt aacttaccct tatctgccaa aaccagagca aaatgctaaa tacgttattg 300
ctaatcagtg gtctcaaatc gatttgcctc cctttgcctc gtctgagggc tgtaagcctg 360
aagatagtgg caagcaccaa gtcagtttcc aaaattgccc ctcagctgct ttaagtgact 420
cagcaccctg cctcagcttc agcaggcgta ggctcaccct gggcggagca aagtatgggc 480
cagggagaac tacagctacg aagacctgct gtcgagttga gaaaagggga gaatttatgg 540
tctgaatttt ctaactgncc tctttnnttg ggtctaaagc tcataataca caaaggcttc 600
cagacctgag cc 612

<210> 293
<211> 440
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 4, 39, 81, 104, 121, 183, 203, 292, 334, 375, 427, 435
<223> n = A,T,C or G

<400> 293
cggnaaggct ggaaaggact ccggaaaggc caagacaang gcggtttccc gctcgcagag 60
agccggcttg cagttcccag ngggccgtat tcatcgacac ctanaatcta ggacgaccag 120
ncatggacgt gtgggcgcga ctgccgctgt gtacagcgca gccatcctgg agtacctcac 180
cgnagaggta cttgaactgg cangaaatgc atcaaaagac ttaaaggtaa agcgtattac 240
ccctcgtcac ttgcaacttg ctattcgtgg agatgaagaa ttggattctc tnatcaaggc 300

tacaattgct ggtggtggtg tcattccaca catncacaaa tctctgattg ggaagaaagg 360
acaacagaag actgnctaaa ggatgcctgg attccttgtt atctcaggac tctaaatact 420
ctaacanctg tccantgttg 440

<210> 294
<211> 423
<212> DNA
<213> Homo sapiens

<400> 294
ctagtccagt gtggtggaat tccttcagta tgatcttgtg ctgtgctatc cgcaggaacc 60
gcgagatggt ctagagtcag cttacatccc tgagcaggaa agtttaccca tgaagattgg 120
tgggattttt tgtttgtttg ttttgttttg tttgttgttt gttgtttgtt tttttgccac 180
taattttagt attcattctg cattgctaga taaaagctga agttacttta tgtttgtctt 240
ttaatgcttc attcaatatt gacatttgta gttgagcggg gggtttggtt tgctttggtt 300
tatatttttt cagttgtttg tttttgcttg ttatattaag cagaaatcct gcaatgaaag 360
gtactatatt tgctagactc tagacaagat attgtacata aaagaatttt tttgtcttta 420
aat 423

<210> 295
<211> 338
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 14, 29, 49, 73, 151, 273
<223> n = A,T,C or G

<400> 295
ctagttagtg cagntttca ttgtgttgng tggttggtct cataactang ttgagttttt 60
ctcctctgct gangaaacag taccgaagtt ctttttcttg tggcatttgt attataaaaa 120
cttggtgtgg gggaggagca caaaactcca ncccactgaa cctctgccaa ttaagatggt 180
gttgggttag gttacatctg gttactgtcc tgggaaaatc atttttatag agatggcctt 240
ccaagtggtt ttaaaattta ctgaagtttt tangtcaatt atgtatgttg actaaattta 300
caaataaact tgtttatcca aaaaaaaaaa aaaagggc 338

<210> 296
<211> 616
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 589, 608
<223> n = A,T,C or G

<400> 296
ctagtccagt gtggtggaat tccgcctcgg aggcgttcag ctgcttcaag atgaagctga 60
acatctcctt cccagccact ggctgccaga aactcattga agtggacgat gaacgcaaac 120
ttcgtacttt ctatgagaag cgtatggcca cagaagttgc tgctgacgct ctgggtgaag 180
aatggaaggg ttatgtggtc cgaatcagtg gtgggaacga caaacaaggt ttccccatga 240
agcagggtgt cttgacccat ggccgtgtcc gcctgctact gagtaagggg cattcctgtt 300
acagaccaag gagaactgga gaaagaaaga gaaaatcagt tcgtggttgc attgtggatg 360
caaatctgag cgttctcaac ttggttattg taaaaaaagg agagaaggat attcctggac 420
tgactgatac tacagtgcct cgccgcctgg gccccaaaag agctagcaga atccgcaaac 480
ttttcaatct ctctaaagaa gatgatgtcc gccagtatgt tgtaagaaag cccttaaata 540
aagaaggtaa gaaacctagg accaaagcac ccaagattca gcgtcttgnt actccacgtg 600

tcctgcanca caaacg 616

<210> 297
<211> 342
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 230, 231
<223> n = A,T,C or G

<400> 297
ctagttagtg cagcttttca ttgtgttgtg tggttggtct cataactagg ttgagttttt 60
ctcctctgct gaggaaacag taccgaagtt ctttttcttg tggcatttgt attataaaaa 120
cttggtgtgg gggaggagca caaaactcca gcccactgaa cctctgccaa ttaagatggt 180
gttgggttag gttacatctg gttactgtcc tgggaaaatc atttttatan nagatggcct 240
tccaagtggt tttaaaattt actgaagttt ttaggtcaat tatgtatgtt gactaaattt 300
acaaataaac ttgtttatcc aaaaaaaaaa aaaaaaaagg gc 342

<210> 298
<211> 456
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 269, 300, 301, 315, 317, 320, 341, 349
<223> n = A,T,C or G

<400> 298
ctagtccagt gtggtggaat tccggagggc cccctcaagg gcatcctggg ctacactgag 60
caccaggtgg tctcctctga cttcaacagc gacacccact cctccacctt cgacgctggg 120
gctggcattg ccctcaacga ccactttgtc aagctcattt cctggtatga caacgaattt 180
ggctacagca acagggtggt ggacctcatg gcccacatgg cctccaagga gtaagacccc 240
tggaccacca gccccagcaa gagcacaana ggaagagaga gaccctcact gctggggagn 300
ncctgccaca ctcantnccn caccacactg aatctcccct nctcacagnt tccatgtaga 360
ccccttgaag aggggagggg cctagggagc cgcaccttgt catgtaccat caataaagta 420
ccctgtgctc aaccaaaaaa aaaaaaaaaa aagggc 456

<210> 299
<211> 570
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 102, 161, 274, 367, 492, 504, 535, 537, 563
<223> n = A,T,C or G

<400> 299
ctagtaggat agaaacactg tgtcccgaga gtaaggagag aagctactat tgattagagc 60
ctaacccagg ttaactgcaa gaagaggcgg gatactttca gntttccatg taactgtatg 120
cataaagcca atgtagtcca gtttctaaga tcatgttcca ngctaactga atcccacttc 180
aatacacact catgaactcc tgatggaaca ataacaggcc caagcctgtg gtatgatgtg 240
cacacttgct agactcagaa aaaatactac tctnataaat gggtgggagt attttggtga 300
caacctactt tgcttggctg agtgaaggaa tgatattcat atattcattt attccatgga 360
catttantta gtgcttttta tataccaggc atgatgctga gtgacactct tgtgtatatt 420

tccaaatttt tgtacagtcg ctgcacatat ttgaaatcat atattaagac tttccaaaga 480
tgaggtccct gnttttttcat ggcnacttga tcagtaagga tttcacctct gtttngnaac 540
taaaaccatc tactatatgt tanacatgac 570

<210> 300
<211> 572
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 562
<223> n = A,T,C or G

<400> 300
ctagtaggat agaaacactg tgtcccgaga gtaaggagag aagctactat tgattagagc 60
ctaacccagg ttaactgcaa gaagaggcgg gatactttca gctttccatg taactgtatg 120
cataaagcca atgtagtcca gtttctaaga tcatgttcca agctaactga atcccacttc 180
aatacacact catgaactcc tgatggaaca ataacaggcc caagcctgtg gtatgatgtg 240
cacacttgct agactcagaa aaaatactac tctcataaat gggtgggagt attttggtga 300
caacctactt tgcttggctg agtgaaggaa tgatattcat atattcattt attccatgga 360
catttagtta gtgcttttta tataccaggc atgatgctga gtgacactct tgtgtatatt 420
tccaaatttt tgtacagtcg ctgcacatat ttgaaatcat atattaagac tttccaaaga 480
tgaggtccct ggtttttcat ggcaacttga tcagtaagga tttcacctct gtttgtaact 540
aaaaccatct actatatgtt angacatgac at 572

<210> 301
<211> 559
<212> DNA
<213> Homo sapiens

<400> 301
ctagtccagt gtggtggaat tccggagccg gcgccctcat gatgctggtg ggcttcctgg 60
gctgctgcgg ggctgtgcag gagtcccagt gcatgctggg actgttcttc ggcttcctct 120
tggtgatatt cgccattgaa atagctgcgg ccatctgggg atattcccac aaggatgagg 180
tgattaagga agtccaggag ttttacaagg acacctacaa caagctgaaa accaaggatg 240
agccccagcg ggaaacgctg aaagccatcc actatgcgtt gaactgctgt ggtttggctg 300
ggggcgtgga acagtttatc tcagacatct gccccaagaa ggacgtactc gaaaccttca 360
ccgtgaagtc ctgtcctgat gccatcaaag aggtcttcga caataaattc cacatcatcg 420
gcgcagtggg catcggcatt gccgtggtca tgatatttgg catgatcttc agtatgatct 480
tgtgctgtgc tatccgcagg aaccgcgaga tggtctagag tcagcttaca tccctgagca 540
ggaaagttta cccatgaag 559

<210> 302
<211> 537
<212> DNA
<213> Homo sapiens

<400> 302
ctagtaggat agaaacactg tgtcccgaga gtaaggagag aagctactat tgattagagc 60
ctaacccagg ttaactgcaa gaagaggcgg gatactttca gctttccatg taactgtatg 120
cataaagcca atgtagtcca gtttctaaga tcatgttcca agctaactga atcccacttc 180
aatacacact catgaactcc tgatggaaca ataacaggcc caagcctgtg gtatgatgtg 240
cacacttgct agactcagaa aaaatactac tctcataaat gggtgggagt attttggtga 300
caacctactt tgcttggctg agtgaaggaa tgatattcat atattcattt attccatgga 360
catttagtta gtgcttttta tataccaggc atgatgctga gtgacactct tgtgtatatt 420
tccaaatttt tgtacagtcg ctgcacatat ttgaaatcat atattaagac tttccaaaga 480
tgaggtccct ggtttttcat ggcaacttga tcagtaagga tttcacctct gtttgta 537

<210> 303
<211> 268
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 23
<223> n = A,T,C or G

<400> 303
ctagttagct ttaagcaccc tanaggacta gggtaatctg acttctcact tcctaagttc 60
ccttctatat cctcaaggta gaaatgtcta tgttttctac tccaattcat aaatctattc 120
ataagtcttt ggtacaagtt tacatgataa aaagaaatgt gatttgtctt cccttctttg 180
cacttttgaa ataaagtatt tatctcctgt ctacagttta ataaatagca tctagtacac 240
aaaaaaaaaa aaaaaaaaaa aaaagggc 268

<210> 304
<211> 434
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 20, 288, 314, 380, 384, 415
<223> n = A,T,C or G

<400> 304
ctagtccagt gtggtggaan tcggagacga cgtgcagaaa tggcacctcg aaaggggaag 60
gaaaagaagg aagaacaggt catcagcctc ggacctcagg tggctgaagg agagaatgta 120
tttggtgtct gccatatctt tgcatccttc aatgacactt ttgtccatgt cactgatctt 180
tctggcaagg aaaccatctg ccgtgtgact ggtgggatga aggtaaaggc agaccgagat 240
gaatcctcac catatgctgc tatgttggct gcccaggatg tggcccanag gtgcaaggag 300
ctgggtatca ccgncctaca catcaaactc cgggccacag gaggaaatag gaccaagacc 360
cctggacctg gggcccagtn cggncctcag agcccttgcc cgctcgggta tgaanatcgg 420
gcggattgag gatg 434

<210> 305
<211> 266
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 20, 38
<223> n = A,T,C or G

<400> 305
ctagtccagt gtggtggaan tcggcgttgg cggcagcntg tggccttcct catctgggcg 60
atgtgggctc ctagaagagt aaggataaca tcctggaaat gacttctgta cggtttgagc 120
ccaactgcac actcatgact tggagctgcc ctgtggagtt acagtttacc aaacacattc 180
atgaacataa tctcatttac taaaaacttt gtgagaattt tcttttacta aaatttttc 240
ttattacaaa aaaaaaaaaa aagggc 266

<210> 306
<211> 236
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<222> 4, 19, 95, 107, 116, 188
<223> n = A,T,C or G

<400> 306
ctantccagt gtggtggant tccgcggcgg tcactgcgcc ggggtagtgg gccccagtgt 60
tgcgctctct ggccgttcct tacactttgc ttcangctcc agtgcanggg cgtagnggga 120
tatggccaac tcgggctgca aggacgtcac gggtccagat gaggagagtt ttctgtactt 180
tgcctacngc agcaacctgc tgacagagag gatccacctc cgaaacccct cggcgg 236

<210> 307
<211> 266
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 257, 262
<223> n = A,T,C or G

<400> 307
ctagtatatg aaaatgtaaa tatcacttgt gtactcaaac aaaagttggt cttaagcttc 60
caccttgagc agccttggaa acctaacctg cctcttttag cataatcaca ttttctaaat 120
gattttcttt gttcctgaaa aagtgatttg tattagtttt acatttgttt tttggaagat 180
tatatttgta tatgtatcat cataaaatat ttaaataaaa agtatcttta gagtgaaaaa 240
aaaaaaaaaa aaaaaanaaa angggc 266

<210> 308
<211> 262
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 20, 21, 23, 39, 94, 142, 155, 170, 185, 187, 204, 214, 215
<223> n = A,T,C or G

<400> 308
ctagtatatg ggtaacaaan nantatgtct gaacctcanc ctataatact ttctactacc 60
tttgcaagga gatgggatag gaacaatcac tcanaggagg cgttgcatgg gcagggtcat 120
agggggaaga aaggtggttt anctgtttta tttanccatt caggggggctn tccatagagg 180
agacngnggt agagggtgaa ctanagaaga taannatgtc ttcctaggcc ggatgcggtg 240
gctcacgcct gtaatcccag ca 262

<210> 309
<211> 419
<212> DNA
<213> Homo sapiens

<400> 309
ctagtgcttt acctttatta atgaactgtg acaggaagcc caaggcagtg ttcctcacca 60
ataacttcag agaagtcagt tggagaaaat gaagaaaaag gctggctgaa aatcactata 120
accatcagtt actggtttca gttgacaaaa tatataatgg tttactgctg tcattgtcca 180
tgcctacaga taatttattt tgtatttttg aataaaaaac atttgtacat tcctgatact 240
gggtacaaga gccatgtacc agtgtactgc tttcaactta aatcactgag gcatttttac 300

tactattctg ttaaaatcag gattttagtg cttgccacca ccagatgaga agttaagcag 360
cctttctgtg gagagtgaga ataattgtgt acaaagtaga gaagtatcca attatgtga 419

<210> 310
<211> 196
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 73
<223> n = A,T,C or G

<400> 310
tgtcatgatt cactattcta gaacttgcat gacctttact gtgttagctc tttgaatgtt 60
cttgaaattt tanactttct ttgtaaacaa atgatatgtc cttatcattg tataaaagct 120
gttatgtgca acagtgtgga gattccttgt ctgatttaat aaaatactta aacactgaaa 180
aaaaaaaaaa aagggc 196

<210> 311
<211> 111
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 8, 43, 101
<223> n = A,T,C or G

<400> 311
tataaaanct tgctgcctga ctaaagatta acaggttata gtntaaattt gtaattaatt 60
ctaccatctt gcaataaagt gacaattgaa tgaaaaaaaa naaaaaaggg c 111

<210> 312
<211> 202
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 13, 33, 40, 71, 99, 129, 195, 196
<223> n = A,T,C or G

<400> 312
aattctaata atnccagctt ctacacagga gtntatattn tgatcggagc cggcgccctc 60
atgatgctgg ngggcttcct gggctgctgc ggggctgtnc aggagtccca gtgcatgctg 120
ggactgttnt tcggcttcct cttggtgata ttcgccattg aaatagctgc ggccatctgg 180
ggatattccc acaanngatg ag 202

<210> 313
<211> 336
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 245, 333
<223> n = A,T,C or G

<400> 313
ctagtctgct gatagaaagc actatacatc ctattgtttc tttctttcca aaatcagcct 60
tctgtctgta acaaaaatgt actttataga gatggaggaa aaggtctaat actacatagc 120
cttaagtgtt tctgtcattg ttcaagtgta ttttctgtaa cagaaacata tttggaatgt 180
ttttcttttc cccttataaa ttgtaattcc tgaaatactg ctgctttaaa aagtcccact 240
gtcanattat attatctaac aattgaatat tgtaaatata cttgtcttac ctctcaataa 300
aagggtactt ttctattaaa aaaaaaaaaa aanggc 336

<210> 314
<211> 315
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 291, 293, 300, 301, 308, 311
<223> n = A,T,C or G

<400> 314
tgcttctgaa ataactctgt attgtagatt atgcagatct ttacaggcat aaatatttaa 60
actgtaatat gctaacttga agagattgca ataaagctgc ttcagctaac cctgtttatg 120
tttaaatact agggtttgtt ctatatttta tacatgcatt ttggatgatt aaagaatgcc 180
tggttttcgt ttgcaatttg cttgtgtaaa tcaggttgta aaaaggcaga taaattgaaa 240
tgtttgtggt atgaggaaat aaaagaatgg aattagcttt caaaaaaaaa nanaaaaaan 300
naaaaaanaa ngggc 315

<210> 315
<211> 277
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 2, 5, 218, 263
<223> n = A,T,C or G

<400> 315
nngtnaagtc aactgcttct gaaataactc tgtattgtag attatgcaga tctttacagg 60
cataaatatt taaactgtaa tatgctaact tgaagagatt gcaataaagc tgcttcagct 120
aaccctgttt atgtttaaat actagggttt gttctatatt ttatacatgc attttggatg 180
attaaagaat gcctggtttt cgtttgcaat ttgcttgngt aaatcaggtt gtaaaaaggc 240
agataaattg aaatgtttgt ggnatgagga aataaaa 277

<210> 316
<211> 599
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 548
<223> n = A,T,C or G

<400> 316
ctagtccagt gtggtggaat tcgcgcggtt gttctctgga gcagcgttct tttatctccg 60
tccgccttct ctcctaccta agtgcgtgcc gccacccgat ggaagattcg atggacatgg 120
acatgagccc cctgaggccc cagaactatc ttttcggttg tgaactaaag gccgacaaag 180

attatcactt taaggtggat aatgatgaaa atgagcacca gttatcttta agaacggtca 240
gtttaggggc tggtgcaaag gatgagttgc acattgttga agcagaggca atgaattacg 300
aaggcagtcc aattaaagta acactggcaa ctttgaaaat gtctgtacag ccaacggttt 360
cccttggggg ctttgaaata acaccaccag tggtcttaag gttgaagtgt ggttcagggc 420
cagtgcatat tagtggacag cacttagtag ctgtggagga agatgcagag tcagaagatg 480
aagaggagga ggatgtgaaa ctcttaagta tatctggaaa gcggtctgcc cctggaggtg 540
gtagcaangt tccacagaaa aaagttaaaa cttgctgctg atgaagatga tgacgatga 599

<210> 317
<211> 573
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 458
<223> n = A,T,C or G

<400> 317
ctagtatatg ggtaacaaat gaatatgtct gaacctcagc tataatactt tctactacct 60
ttgcaaggag atgggatagg aacaatcact cagaggaggc gttgcatggg cagggtcata 120
gggggaagaa aggtggttta gctgttttat ttagccattc agggggctct ccagagagga 180
gacggtggta gagggtgaac tagagaagat aagaatgtct tcctaggccg gatgcggtgg 240
ctcacgcctg taatcccagc actttgggat tgcgaggtgg gcggatcact tgaggtcagg 300
agttcaagac cagcctggcc aacatggtaa aacccgtctc tactaacaat acaaagatta 360
gcctggtgtg gtggcacggg cctgtaatcg cagccccttg gaaggccaag gcaggagaat 420
cgcctcaaca ctggaggtgg aggttgcagt gagctganat tgtgccactg cactccagcc 480
tgggcaatga ggcaagaccc tgtctcaaaa aataataaat aataataata ataatgtttt 540
tctagagttt cagtctaagg gaaaatgtga ttt 573

<210> 318
<211> 547
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 4, 5
<223> n = A,T,C or G

<400> 318
ctannccagt gtggtggaat tcgcgccagg tcccgccagt cccagctgcg cgcgcccccc 60
agtcccgcac ccgttcggcc caggctaagt tagccctcac catgccggtc aaaggaggca 120
ccaagtgcat caaatacctg ctgttcggat ttaacttcat cttctggctt gccgggattg 180
ctgtccttgc cattggacta tggctccgat tcgactctca gaccaagagc atcttcgagc 240
aagaaactaa taataataat tccagcttct acacaggagt ctatattctg atcggagccg 300
gcgccctcat gatgctggtg ggcttcctgg gctgctgcgg ggctgtgcag gagtcccagt 360
gcatgctggg actgttcttc ggcttcctct tggtgatatt cgccattgaa atagctgcgg 420
ccatctgggg atattcccac aaggatgagg tgattaagga agtccaggag ttttacaagg 480
acacctacaa caagctgaaa accaaggatg agccccagcg ggaaacgctg aaagccatcc 540
actatgc 547

<210> 319
<211> 529
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<222> 6, 251
<223> n = A,T,C or G

<400> 319
ctagtncagt gtggtggaat tcgaagaacc atgggtggac ccgaactccc cggtgctctt 60
ggaggaccca gtcctttgtg ccttggcaaa aaagcacaag cgaaccccag ccctgattgc 120
cctgcgctac cagctgcagc gtggggttgt ggtcctggcc aagagctaca atgagcagcg 180
catcagacag aacgtgcagg tttttgagtt ccagttgact gcagaggaca tgaaagccat 240
agatggccta nacagaaatc tccactattt taacagtgat agttttgcta gccaccctaa 300
ttatccatat tcagatgaat attaacatgg agagctttgc ctgatgtcta ccagaagccc 360
tgtgtgtgga tggtgacgca gaggacgtct ctatgccggt gactggacat atcacctcta 420
cttaaatccg tcctgtttag cgacttcagt caactacagc tgagtccata ggccaggaaa 480
gacaataaat ttttatcatt ttgaaataaa aaaaaaaaaa aaaaagggc 529

<210> 320
<211> 225
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 163
<223> n = A,T,C or G

<400> 320
ctagtccagt gtggnggaat tctaataatt ccagcttcta cacaggagtc tatattctga 60
tcggagccgg cgccctcatg atgctggtgg gcttcctggg ctgctgcggg gctgtgcagg 120
agtcccagtg catgctggga ctgttcttcg gcttcctctt ggngatattc gccattgaaa 180
tagctgcggc catctgggga tattcccaca aggatgaggt gatta 225

<210> 321
<211> 308
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 6, 13, 15, 50, 220, 236, 247, 262, 281, 287, 299, 302
<223> n = A,T,C or G

<400> 321
ctagtncagt gtngnggaat tctaataatt ccagcttcta cacaggagtn tatattctga 60
tcggagccgg cgccctcatg atgctggtgg gcttcctggg ctgctgcggg gctgtgcagg 120
agtcccagtg catgctggga ctgttcttcg gcttcctctt ggtgatattc gccattgaaa 180
tagctgcggc catctgggga tattcccaca aggatgaggn gattaaggaa gtccangagt 240
tttacangga cacctacaac angctgaaaa ccaaggatga nccccancgg gaaacgctna 300
angccatc 308

<210> 322
<211> 567
<212> DNA
<213> Homo sapiens

<400> 322
ctagtccagt gtggtggaat tcgtgtcttt tcactaatta cctatactat gccaatattt 60
ccttatatct atccataaca tttatactac atttgtaaga gaatatgcac gtgaaactta 120
acactttata aggtaaaaat gaggtttcca agatttaata atctgatcaa gttcttgtta 180

tttccaaata gaatggactt ggtctgttaa gggctaagga gaagaggaag ataaggttaa 240
aagttgttaa tgaccaaaca ttctaaaaga aatgcaaaaa aaaagtttat tttcaagcct 300
tcgaactatt taaggaaagc aaaatcattt cctaaatgca tatcatttgt gagaatttct 360
cattaatatc ctgaatcatt catttcagct aaggcttcat gttgactcga tatgtcatct 420
aggaaagtac tatttcatgg tccaaacctg ttgccatagt tggtaaggct ttcctttaag 480
ttgtgaaata tttagatgaa attttctctt ttaaagttct ttatagggtt agggtgtggg 540
aaaatgctat attaataaat ctgtagt 567

<210> 323
<211> 598
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15
<223> n = A,T,C or G

<400> 323
ctagtccagt gtggnggaat tccttcgcct tagtactcgt gtgaagttgg cggggacggt 60
tcctgtcatc ttcttgggct tattttgtgt gctgttgaag gggggagact agagaaatgg 120
cagggaacct cttatccggg gcaggtaggc gcctgtggga ctgggtgcct ctggcgtgca 180
gaagcttctc tcttggtgtg cctagattga tcggtataag gctcactctc ccgcccccca 240
aagtggttga tcgttggaac gagaaaaggg ccatgttcgg agtgtatgac aacatcggga 300
tcctgggaaa ctttgaaaag caccccaaag aactgatcag ggggcccata tggcttcgag 360
gttggaaagg gaatgaattg caacgttgta tccgaaagag gaaaatggtt ggaagtagaa 420
tgttcgctga tgacctgcac aaccttaata aacgcatccg ctatctctac aaacacttta 480
accgacatgg gaagtttcga tagaagagaa agctgagaac ttcggaaaag gctcatctgt 540
caccctggag aagggaaact gtacttttcc ctgtgaggaa acggctttgt attttctc 598

<210> 324
<211> 223
<212> DNA
<213> Homo sapiens

<400> 324
ctagtgagct ctaggctgta gaaatttaaa aactacaatg tgattaactc gagcctttag 60
ttttcatcca tgtacatgga tcacagtttg ctttgatctt cttcaatatg tgaatttggg 120
ctcacagaat caaagcctat gcttggttta atgcttgcaa tctgagctct tgaacaaata 180
aaattaacta ttgtagtgtg aaaaaaaaaa aaaaaaaaag ggc 223

<210> 325
<211> 500
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 338, 339, 348, 356, 374, 383, 410, 451, 469, 490
<223> n = A,T,C or G

<400> 325
ggaattctaa taattccagc ttctacacag gagtctatat tctgatcgga gccggcgccc 60
tcatgatgct ggtgggcttc ctgggctgct gcggggctgt gcaggagtcc cagtgcatgc 120
tgggactgtt cttcggcttc ctcttggtga tattcgccat tgaaatagct gcggccatct 180
ggggatattc ccacaaggat gaggtgatta aggaagtcca ggagttttac aaggacacct 240
acaacaagct gaaaaccaag gatgagcccc agcgggaaac gctgaaagcc atccactatg 300
cgttgaactg ctgtggtttg gctgggggcg tggaacannt tatctcanac atctgnccca 360

agaaggacgt actngaaacc ttnaccgtga agtcctgtcc tgatgccatn aaagaggtct 420
tcgacaataa attccacatc atcggcgcag ngggcatcgg cattgccgng gtcatgatat 480
ttggcatgan cttcagtatg 500

<210> 326
<211> 515
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 292, 322, 325, 356, 380, 383, 418, 420, 476, 479, 484, 500, 504, 506
<223> n = A,T,C or G

<400> 326
agtgtggtgg aattctaata attccagctt ctacacagga gtctatattc tgatcggagc 60
cggcgccctc atgatgctgg tgggcttcct gggctgctgc ggggctgtgc aggagtccca 120
gtgcatgctg ggactgttct tcggcttcct cttggtgata ttcgccattg aaatagctgc 180
ggccatctgg ggatattccc acaaggatga ggtgattaag gaagtccagg agttttacaa 240
ggacacctac aacaagctga aaaccaagga tgagccccag cgggaaacgc tnaaagccat 300
ccactatgcg ttgaactgct gnggnttggc tgggggcgtg gaacagttta tctcanacat 360
cctgccccaa gaaggacgtn ctngaaacct tcaccgttga agtcctgtcc tgatgccntn 420
aaagaggtct tcgacaataa attccacatc atcggcgcag tgggcatcgg cattgncgng 480
gtcntgatat ttggcatgan cttnantatg atctt 515

<210> 327
<211> 466
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 339, 348, 374, 383, 451
<223> n = A,T,C or G

<400> 327
ggaattctaa taattccagc ttctacacag gagtctatat tctgatcgga gccggcgccc 60
tcatgatgct ggtgggcttc ctgggctgct gcggggctgt gcaggagtcc cagtgcatgc 120
tgggactgtt cttcggcttc ctcttggtga tattcgccat tgaaatagct gcggccatct 180
ggggatattc ccacaaggat gaggtgatta aggaagtcca ggagttttac aaggacacct 240
acaacaagct gaaaaccaag gatgagcccc agcgggaaac gctgaaagcc atccactatg 300
cgttgaactg ctgtggtttg gctgggggcg tggaacagnt tatctcanac atctgcccca 360
agaaggacgt actngaaacc ttnaccgtga agtcctgtcc tgatgccatc aaagaggtct 420
tcgacaataa attccacatc atcggcgcag ngggcatcgg cattgc 466

<210> 328
<211> 481
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 220, 329, 332, 356, 413, 438
<223> n = A,T,C or G

<400> 328
ctagtccagt gtggnggaat tctaataatt ccagcttcta cacaggagtc tatattctga 60

tcggagccgg cgccctcatg atgctggtgg gcttcctggg ctgctgcggg gctgtgcagg 120
agtcccagtg catgctggga ctgttcttcg gcttcctctt ggtgatattc gccattgaaa 180
tagctgcggc catctgggga tattcccaca aggatgaggn gattaaggaa gtccaggagt 240
tttacaagga cacctacaac aagctgaaaa ccaaggatga gccccagcgg gaaacgctga 300
aagccatcca ctatgcgttg aactgctgng gnttggctgg gggcgtggaa cagttnatct 360
cagacatctg ccccaagaag gacgtactcg aaaccttcac cgtgaagtcc tgncctgatg 420
ccatcaaaga ggtcttcnga caataaattc cacatcatcg gcgcagtggg catcggcatt 480
g 481

<210> 329
<211> 355
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 15, 50, 155, 189, 237, 263, 282, 300, 316, 318, 333
<223> n = A,T,C or G

<400> 329
ctagtccagt gtggnggaat tctaataatt ccagcttcta cacaggagtn tatattctga 60
tcggagccgg cgccctcatg atgctggtgg gcttcctggg ctgctgcggg gctgtgcagg 120
agtcccagtg catgctggga ctgttcttcg gcttnctctt ggtgatattc gccattgaaa 180
tagctgcang ccatctgggg atattcccac aaggatgagg tgattaagga agtccangag 240
ttttacaagg acacctacaa cangctgaaa accaaggatg ancccagcg ggaaacgctn 300
aaagccatcc actatncntt gaactgctgt ggnttggctg ggggcgtgga acagt 355

<210> 330
<211> 179
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 20, 49, 91, 120, 155, 157, 160
<223> n = A,T,C or G

<400> 330
cctggtcttg agatgtcttn tcgttaagga gatgggcctt ttggaggtna aggataaaat 60
gaatgagttc tgtcatgatt cactattcta naacttgcat gacctttact gtgttagctn 120
tttgaatgtt cttgaaattt tagactttct ttgtnancan ataatatgtc cttatcatt 179

<210> 331
<211> 565
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 420, 455, 504, 505, 559
<223> n = A,T,C or G

<400> 331
ctagttagtt ctactaatta gaaacttgct gtacttttc ttttctttta ggggtcaagg 60
accctcttta tagctaccat ttgcctacaa taaattattg cagcagtttg caatactaaa 120
atatttttta tagactttat atttttcctt ttgataaagg gatgctgcat agtagagttg 180
gtgtaattaa actatctcag ccgtttccct gctttccctt ctgctccata tgcctcattg 240
tccttccagg gagctctttt aatcttaaag ttctacattt catgctctta gtcaaattct 300

gttacctttt taataactct tcccactgca tatttccatc ttgaattggt ggttctaaat 360
tctgaaactg tagttgagat acagctattt aatatttctg ggagatgtgc atccctcttn 420
tttgtggttg cccaaggttg ttttgcgtaa ctganactcc ttgatatgct tcagagaatt 480
taggcaaaca ctggccatgg ccgnngggag tactgggagt aaaataaaaa tatcgaggta 540
tagactagca tccacatana gcact 565

<210> 332
<211> 476
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 415
<223> n = A,T,C or G

<400> 332
ctagtgagga cgttaaccag ccatattggc tcaataaata gcttcggtaa ggagttaatt 60
tccttctaga aatcagtgcc tatttttcct ggaaactcaa ttttaaatag tccaattcca 120
tctgaagcca agctgttgtc attttcattc ggtgacattc tctcccatga cacccagaag 180
gggcagaaga accacatttt tcatttatag atgtttgcat cctttgtatt aaaattattt 240
tgaaggggtt gcctcattgg atggcttttt ttttttcct ccagggagaa ggggagaaat 300
gtacttggaa attaatgtat gtttacatct ctttgcaaat tcctgtacat agagatatat 360
tttttaagtg tgaatgtaac aacatactgt gaattccatc ttggttacaa atganactcc 420
ttcagtcagt tatccaaata aaagcagttc tgaaactaaa aaaaaaaaaa aaaagg 476

<210> 333
<211> 458
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 450
<223> n = A,T,C or G

<400> 333
ctagtccagt gtggtggaat tctggagacg acgtgcagaa atggcacctc gaaaggggaa 60
ggaaaagaag gaagaacagg tcatcagcct cggacctcag gtggctgaag gagagaatgt 120
atttggtgtc tgccatatct ttgcatcctt caatgacact tttgtccatg tcactgatct 180
ttctggcaag gaaaccatct gccgtgtgac tggtgggatg aaggtaaagg cagaccgaga 240
tgaatcctca ccatatgctg ctatgttggc tgcccaggat gtggcccaga ggtgcaagga 300
gctgggtatc accgccctac acatcaaact ccgggccaca ggaggaaata ggaccaagac 360
ccctggacct ggggcccagt cggccctcag agcccttgcc cgctcgggta tgaagatcgg 420
gcggattgag gatgtcaccc ccatcccctn tgacagca 458

<210> 334
<211> 568
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 523, 529, 534
<223> n = A,T,C or G

<400> 334
ctagtccagt gtggtggaat tcgaacagta ttgctgtaat tccttttctt ttcttcctca 60

tttcctctgc cccttaaaag attgaagaaa gagaaacttg tcaactcata tccacgttat 120
ctagcaaagt acataagaat ctatcactaa gtaatgtatc cttcagaatg tgttggttta 180
ccagtgacac cccatattca tcacaaaatt aaagcaagaa gtccatagta atttatttgc 240
taatagtgga tttttaatgc tcagagtttc tgaggtcaaa ttttatcttt tcacttacaa 300
gctctatgat cttaaataat ttacttaatg tattttggtg tattttcctc aaattaatat 360
tggtgttcaa gactatatct aattcctctg atcactttga gaaacaaact tttattaaat 420
gtaaggcact tttctatgaa ttttaaatat aaaaataaat attgttctga ttattactga 480
aaagatgtca gccatttcaa tgtcttggga aacaattttt tgnttttgnt ctgntttctt 540
tttgcttcaa taaaacaata gctggctc 568

<210> 335
<211> 450
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 26, 43, 176, 180, 213, 229, 232, 255, 274, 322, 325, 373,
382, 391, 396, 419, 430, 431
<223> n = A,T,C or G

<400> 335
agtgtggtgg aattctaata attccngctt ctacacagga gtntatattc tgatcggagc 60
cggcgccctc atgatgctgg tgggcttcct gggctgctgc ggggctgtgc aggagtccca 120
gtgcatgctg ggactgttct tcggcttcct cttggtgata ttcgccattg aaatanctgn 180
ggccatctgg ggatattccc acaaggatga ggngattaag gaagtccang anttttacaa 240
ggacacctac aacangctga aaaccaagga tgancccag cgggaaacgc tgaaagccat 300
ccactatgcg ttgaactgct gnggnttggc tgggggcgtg gaacagttta tctcagacat 360
ctgccccaag aangacgtac tngaaacctt naccgngaag tcctgtcctg atgccatcna 420
agaggtcttn nacaataaat tccacatcat 450

<210> 336
<211> 555
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 45, 129, 160, 220, 262, 281, 329, 356, 371, 389, 459, 465,
478, 484, 511
<223> n = A,T,C or G

<400> 336
ctagtccagt gtggtggaat tctaataatt ccagcttcta cacangagtc tatattctga 60
tcggagccgg cgccctcatg atgctggtgg gcttcctggg ctgctgcggg gctgtgcagg 120
agtcccagng catgctggga ctgttcttcg gcttcctctn ggtgatattc gccattgaaa 180
tagctgcggc catctgggga tattcccaca aggatgaggn gattaaggaa gtccaggagt 240
tttacaagga cacctacaac angctgaaaa ccaaggatga nccccagcgg gaaacgctga 300
aagccatcca ctatgcgttg aactgctgng gtttggctgg gggcgtggaa cagttnatct 360
cagacatctg ncccaagaag gacgtactng aaaccttcac cgtgaagtcc tgtcctgatg 420
ccatcaaaga ggtcttcgac aataaattcc acatcatcng cgcantgggc atcggcantg 480
ccgnggtcat gatatttggc atgatcttca ntatgatctt gtgctgtgct atccgcagga 540
accgcgagat ggtct 555

<210> 337
<211> 368
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 6, 30, 33, 88, 144, 167, 187, 212, 218, 237, 239, 244, 262,
281, 299, 315, 323, 329, 332, 354, 356
<223> n = A,T,C or G

<400> 337
ctagtncagt gtggtggaat tctaataatn ccngcttcta cacaggagtc tatattctga 60
tcggagccgg cgccctcatg atgctggngg gcttcctggg ctgctgcggg gctgtgcagg 120
agtcccagtg catgctggga ctgntcttcg gcttcctctt ggtgatnttc gccattgaaa 180
tagctgnggc catctgggga tattcccaca angatgangt gattaaggaa gtccagnant 240
tttncaagga cacctacaac angctgaaaa ccaaggatga nccccagcgg gaaacgctna 300
aagccatcca ctatncgttg aantgctgng gnttggctgg gggcgtggaa cagntnatct 360
cagacatc 368

<210> 338
<211> 320
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 27, 44, 101, 152, 165, 198, 202, 214, 230, 233, 256, 275,
279, 283, 293, 311, 312
<223> n = A,T,C or G

<400> 338
cagtgtggtg gaattctaat aattccngct tctacacagg agtntatatt ctgatcggag 60
ccggcgccct catgatgctg gtgggcttcc tgggctgctg nggggctgtg caggagtccc 120
agtgcatgct gggactgttc ttcggcttcc tnttggtgat attcnccatt gaaatagctg 180
cggccatctg gggatatncc cncaaggatg aggngattaa ggaagtccan gantttaca 240
aggacaccta caacangctg aaaaccaagg atganncccna gcnggaaacg ctnaaagcca 300
tccactatgc nntgaactgc 320

<210> 339
<211> 599
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 462, 463, 489, 508, 568, 574
<223> n = A,T,C or G

<400> 339
ctagtcacta ctgtcttctc cttgtagcta atcaatcaat attcttccct tgcctgtggg 60
cagtggagag tgctgctggg tgtacgctgc acctgcccac tgagttgggg aaagaggata 120
atcagtgagc actgttctgc tcagagctcc tgatctaccc caccccctag gatccaggac 180
tgggtcaaag ctgcatgaaa ccaggccctg gcagcaacct gggaatggct ggaggtggga 240
gagaacctga cttctctttc cctctccctc ctccaacatt actggaactc tatcctgtta 300
ggatcttctg agcttgtttc cctgctgggt gggacagagg acaaaggaga agggagggtc 360
tagaagaggc agcccttctt tgtcctctgg ggtaaatgag cttgacctag agtaaatgga 420
gagaccaaaa gcctctgatt tttaatttcc ataaaatgtt annaagtata tatatacata 480
tatatattnt ctttaaattt ttgagtcntt tgatatgtct aaaaatccat tccctctgcc 540
ctgaagcctg agtgagacac atgaaganaa ctgngtttca tttaaagatg ttaattaaa 599

<210> 340

<211> 594
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 6, 262, 484, 533, 558, 583
<223> n = A,T,C or G

<400> 340
ctagtncagt gtggtggaat tctaataatt ccagcttcta cacaggagtc tatattctga 60
tcggagccgg cgccctcatg atgctggtgg gcttcctggg ctgctgcggg gctgtgcagg 120
agtcccagtg catgctggga ctgttcttcg gcttcctctt ggtgatattc gccattgaaa 180
tagctgcggc catctgggga tattcccaca aggatgaggt gattaaggaa gtccaggagt 240
tttacaagga cacctacaac angctgaaaa ccaaggatga gccccagcgg gaaacgctga 300
aagccatcca ctatgcgttg aactgctgtg gtttggctgg gggcgtggaa cagtttatct 360
cagacatctg ccccaagaag gacgtactcg aaaccttcac cgtgaagtcc tgtcctgatg 420
ccatcaaaga ggtcttcgac aataaattcc acatcatcgg cgcagtgggc atcggcattg 480
ccgnggtcat gatatttggc atgatcttcc agtatgatct tgtgctgtgc tanccgcagg 540
aaccgcgaga tggtctanag tcagcttaca tccctgagca ggnaagttta ccca 594

<210> 341
<211> 327
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 30, 33, 45, 50, 71, 72, 88, 122, 144, 145, 150, 158, 160,
169, 183, 187, 204, 212, 218, 220, 224, 236, 239, 247, 262,
281, 299, 306, 317, 323
<223> n = A,T,C or G

<400> 341
ctagtccagt gtggtggaat tctaataatn ccngcttcta cacangagtn tatattctga 60
tcggagccgg nnccctcatg atgctggngg gcttcctggg ctgctgcggg gctgtgcagg 120
antcccagtg catgctggga ctgnncttcn gcttcctntn ggtgatatnc gccattgaaa 180
tanctgnggc catctgggga tatncccaca angatgangn gatnaaggaa gtccangant 240
tttacangga cacctacaac angctgaaaa ccaaggatga nccccagcgg gaaacgctna 300
aagccntcca ctatgcnttg aantgct 327

<210> 342
<211> 601
<212> DNA
<213> Homo sapiens

<400> 342
ctagtccagt gtggtggaat tcggcgtgca ggagtcagag acattacatc aggaagatac 60
tgcagagata ttctactcca tctcattcat tgtacagatt ctaaactccc tgaaggagac 120
aaattaccag tggacaagaa cacagcctct ggagtccaat aggcctggtg tattcattag 180
ggatgcctaa atcaaaggaa cttgtttctt caagctcttc tggcagtgat tctgacagtg 240
aggttgacaa aaagttaaag aggaaaaagc aagttgctcc agaaaaacct gtaaagaaac 300
aaaagacagg tgagacttcg agagccctgt catcttctaa acagagcagc agcagcagag 360
atgataacat gtttcagatt gggaaaatga ggtacgttag tgttcgcgat tttaaaggca 420
aagtgctaat tgatattaga gaatattgga tggatcctga aggtgaaatg aaaccaggaa 480
gaaaaggtat ttctttaaat ccagaacaat ggagccagct gaaggaacag atttctgaca 540
ttgatgatgc agtaagaaaa ctgtaaaatt cgagccatat aaataaaacc tgtactgttc 600
t 601

<210> 343
<211> 601
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 99, 143, 148, 168, 183, 224, 228, 229, 278, 304, 346, 348, 363, 516, 517, 519, 550, 573, 582, 589
<223> n = A,T,C or G

<400> 343
ctagtccagt gtggtggaat tcctcccccc gagcgccgct ccggctgcac cgcgctcgct 60
ccgagtttca ggctcgtgct aagctagcgc cgtcgtcgnc tcccttcagt cgccatcatg 120
attatctacc gggacctcat canccacnat gagatgttct ccgacatnta caagatccgg 180
ganatcgcgg acgggttgtg cctggaggtg gaggggaaga tggncagnng gacagaaggt 240
aacattgatg actcgctcat tggtggaaat gcctccgntg aaggccccga gggcgaaggt 300
accnaaagca cagtaatcac tggtgtcgat attgtcatga accatnanct gcaggaaaca 360
agnttcacaa aagaagccta caagaagtac atcaaagatt acatgaaatc aatcaaaggg 420
aaacttgaag aacagagacc agaaagagta aaacctttta tgacaggggc tgcagaacaa 480
atcaagcaca tccttgctaa tttcaaaaac taccanntnt ttattggtga aaacatgaat 540
ccagatggcn tggttgctct attggactac cgngaggatg gngtgaccnc atatatgatt 600
t 601

<210> 344
<211> 388
<212> DNA
<213> Homo sapiens

<400> 344
ctagtccagt gtggtggaat tcatctatac tagataatcc tagatgaaat gttagagatg 60
ctatttgata caactgtggc catgactgag gaaaggagct cacgcccaga gactgggctg 120
ctctcccgga ggccaaaccc aagaaggtct ggcaaagtca ggctcaggga gactctgccc 180
tgctgcagac ctcggtgtgg acacacgctg catagagctc tccttgaaaa cagaggggtc 240
tcaagacatt ctgcctacct attagctttt ctttattttt ttaacttttt gggggaaaa 300
gtatttttga gaagtttgtc ttgcaatgta tttataaata gtaaataaag tttttaccat 360
taaaaaaata aaaaaaaaaa aaaagggc 388

<210> 345
<211> 602
<212> DNA
<213> Homo sapiens

<400> 345
ctagtgatca gtggtcgtga agtgtttgaa tttcgtcctg aactggtcaa tgatgatgat 60
gaggaagcag atgatacccg ctacacccag ggaacaggtg gtgatgaggt tgatgattca 120
gtgagtgtaa atgacataga tttaagcctg tacatcccaa gagatgtaga tgaaacaggt 180
attactgtag ccagtcttga aagattcagc acatatactt cagataaaga tgaaaacaaa 240
ttaagtgaag cttctggagg tagggctgaa aatggtgaaa gaagtgactt ggaagaggac 300
aacgagaggg agggaacgga aaatggagcc attgatgctg ttcctgttga tgaaaatctt 360
ttcactggag aggatttgga tgaactagaa gaagaattaa atacacttga tttagaagaa 420
tgacaccaaa cacatcgctg aaaaaattaa gtcagctcag cacgagttga aattgactac 480
attaatttct ttccacctag aatcaacagg atgtttattt cctatgctga ttctggagga 540
gttaacctcc tgcaaaaaag gcatcttgtc cctacatctt ctcttctgac tttggctaca 600
tc 602

<210> 346

<211> 600
<212> DNA
<213> Homo sapiens

<400> 346
ctagtgactg agttcctggc aaagaaattt gacctggacc agttgataac tcatgtttta 60
ccatttaaaa aaatcagtga aggatttgag ctgctcaatt caggacaaag cattcgaacg 120
gtcctgacgt tttgagatcc aaagtggcag gaggtctgtg ttgtcatggt gaactggagt 180
ttctcttgtg agagttccct catctgaaat catgtatctg tctcacaaat acaagcataa 240
gtagaagatt tgttgaagac atagaaccct tataaagaat tattaacctt tataaacatt 300
taaagtcttg tgagcacctg ggaattagta taataacaat gttaatattt ttgatttaca 360
ttttgtaagg ctataattgt atcttttaag aaaacataca cttggatttc tatgttgaaa 420
tggagatttt taagagtttt aaccagctgc tgcagatata tatctcaaaa cagatatagc 480
gtataaagat atagtaaatg catctcctag agtaatattc acttaacaca ttgaaactat 540
tatttttag atttgaatat aaatgtattt tttaaacact tgttatgagt taacttggat 600

<210> 347
<211> 57
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 3, 4, 6, 16
<223> n = A,T,C or G

<400> 347
ctnnanggca cagtcnaggc tgatcagcgg gtttaaacgg gccctctaga ctcgagc 57

<210> 348
<211> 596
<212> DNA
<213> Homo sapiens

<400> 348
ctagtttatt tccttaaata ttgctacaaa aggaagatgc gggtgtaagc cctgattttt 60
ttttctccca agaaaaatct taaaggacca ctttagataa tatttgattc ctactgtaaa 120
atttagaaaa tgatgaattc ttgtccattt ttgtaatcaa gattttagga aaaacagaag 180
tacatctatc tttatgaaat tttgggcagg tttttgtgta tcaatatttt gtacttttag 240
ggaatatttt atttttagt tatttgtgtc aaattataat tataaaaggt acagcagaaa 300
atataccatg tttttatata ggttcacacc tgtacttagg agggaccctg tccatctata 360
tacttttttgt ataaaatttt aaaatgttaa agatccacaa ggtcttaata aaatgattct 420
atagctagaa aaacatttac cttcccagtg ctttgcacta aaatatactg tgaaaggaaa 480
ctagaaagac tgtaactatt gctggaaatg ttctatattg aatgtacatg ctcttgttgg 540
aaaaatgtac tatatgtgat ggaaataaac cagaatcgaa gttatttcag ctaaat 596

<210> 349
<211> 571
<212> DNA
<213> Homo sapiens

<400> 349
ctagtccagt gtggtggaat tcgcgcagac cagacttcgc tcgtactcgt gcgcctcgct 60
tcgcttttcc tccgcaacca tgtctgacaa acccgatatg gctgagatcg agaaattcga 120
taagtcgaaa ctgaagaaga cagagacgca agagaaaaat ccactgcctt ccaaagaaac 180
gattgaacag gagaagcaag caggcgaatc gtaatgaggc gtgcgccgcc aatatgcact 240
gtacattcca caagcattgc cttcttattt tacttctttt agctgtttaa ctttgtaaga 300

tgcaaagagg ttggatcaag tttaaatgac tgtgctgccc ctttcacatc aaagaactac 360
tgacaacgaa ggccgcgcct gcctttccca tctgtctatc tatctggctg gcagggaagg 420
aaagaacttg catgttggtg aaggaagaag tggggtggaa gaagtggggt gggacgacag 480
tgaaatctag agtaaaacca agctggccca aggtgtcctg caggctgtaa tgcagtttaa 540
tcagagtgcc attttttttt ttgttcaaat g 571

<210> 350
<211> 601
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 549, 553, 561
<223> n = A,T,C or G

<400> 350
ctagtgaatg aagaacgaac gctggaagta gaaatagagc ctggggtgag agacggcatg 60
gagtacccct ttattggaga aggtgagcct cacgtggatg gggagcctgg agatttacgg 120
ttccgaatca aagttgtcaa gcacccaata tttgaaagga gaggagatga tttgtacaca 180
aatgtgacaa tctcattagt tgagtcactg gttggctttg agatggatat tactcacttg 240
gatggtcaca aggtacatat ttcccgggat aagatcacca ggccaggagc gaagctatgg 300
aagaaagggg aagggctccc caactttgac aacaacaata tcaagggctc tttgataatc 360
acttttgatg tggattttcc aaaagaacag ttaacagagg aagcgagaga aggtatcaaa 420
cagctactga aacaagggtc agtgcagaag gtatacaatg gactgcaagg atattgagag 480
tgaataaaat tggactttgt ttaaaataag tgaataagcg atatttatta tctgcaaggg 540
ttttttgng tgngtttttg nttttatttt caatatgcaa gttaggctta atttttttat 600
c 601

<210> 351
<211> 501
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 388, 397
<223> n = A,T,C or G

<400> 351
ctagtccagt gtggttgaat tcccgagctg gaggagctgg gtgtggggtg cgttgggctg 60
gtggggaggc ctagtttggg tgcaagtagg tctgattgag cttgtgttgt gctgaaggga 120
cagccctggg tctaggggag agagtccctg agtgtgagac ccgccttccc cggtcccagc 180
ccctcccagt tccccccaggg acggccactt cctggtcccc gacgcaacca tggctgaaga 240
acaaccgcag gtcgaattgt tcgtgaaggc tggcagtgat ggggccaaga ttgggaactg 300
cccattctcc cagagactgt tcatggtact gtggctcaag ggagtcacct tcaatgttac 360
caccgttgac accaaaaggc ggaccganac agtgcanaag ctgtgcccag gggggcagct 420
cccattcctg ctgtatggca ctgaagtgca cacagacacc aacaagattg aggaatttct 480
ggaggcagtg ctgtgccctc c 501

<210> 352
<211> 475
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 359, 445

<223> n = A,T,C or G

<400> 352
ctagtccagt gtggtggaat tcgccggccc ccagcccgga agttatgaga tccgacacta 60
tggaccagcc aagtgggtca gcacgtccgt ggagtctatg gactgggatt cagccatcca 120
gacgggcttt acgaaactga acagctacat tcaaggcaaa aacgagaaag agatgaaaat 180
aaagatgaca gctccagtga caagctacgt ggagcctggt tcaggtcctt ttagtgagtc 240
taccattacc atttccctgt atattccctc tgaacagcaa tttgatccac ccaggccttt 300
agagtcagat gtcttcattg aagatagagc cgaaatgact gtgtttgtac ggtctttcna 360
tggattttct agtgcccaaa agaatcaaga acaacttttg acattagcaa gcattttaag 420
ggaagatgga aaagttttcg atganaaggt ttactacact gcaggctaca acagt 475

<210> 353
<211> 336
<212> DNA
<213> Homo sapiens

<400> 353
ctagtccatg ccaggacacc agctgacaat ttcttggttt tactgtcaat aattgtacca 60
tgtgatcaat tactgtcctc acttagaaca aagcctgagt ccgagaatat ttatatttta 120
ccaatatatg cctgttacaa gagaaggaaa tatgagttat ttaagtttaa ctttttatg 180
tgaattcaga gtttatttat cgagggaaat atgtacaaag aagcttcaaa tggaatattt 240
accgacattc cttatacatg acagacactt ggctacatgg gaagatgatg ttaataataa 300
aatgattttt aaatggaaaa aaaaaaaaaa aagggc 336

<210> 354
<211> 362
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 314, 361
<223> n = A,T,C or G

<400> 354
ctagtccagt gtggtggaat tctttaaatc tggtccaaag tctttaaaat aggtagattt 60
tcagctttct taagtttctc cctcatttag atttcatggt ttttacataa agggtgaata 120
tttgaatttt cttttaaatt tcactgcatc ttcaattgcc caactgtgtt tcctgataaa 180
ttttagattc acatttttag gaaatttgga gtattccaga caatatacta gatacccaga 240
aacttttctc agtaggttct gaggtgtttt aagttcttat gctagactgt aagctccttg 300
agggcagaga ctgntttatt tattcttgta tcctcagtgc ctggtacagg acttgacaca 360
na 362

<210> 355
<211> 398
<212> DNA
<213> Homo sapiens

<400> 355
ctagtgcttc tggcgatgac atttctaagc tacagcgtac tccaggagaa gaaaagatta 60
ataccttaaa agaagaaaac actcaagaag cagcagtcct gaatggtgtt tcataaactg 120
aagaagttcc tagtttacag ttcttttaca ttacatttac aatagtgctt gtacaagctt 180
gccaaagata gaatatggat cgccagtctt tacatcgcac tttcagttcc tccatttgga 240
attcaaaaag gggagggatc ctgaagaaat catatgttaa acatactttg acacctactg 300
tgttataaaa tatatcatca gatgtgcctt gagaatagta tatgtaacat taaaaaaaag 360
ttgctggcta taggaaaaaa aaaaaaaaaa aaaggggc 398

<210> 356
<211> 144
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 6, 12, 14, 57, 80, 88, 103, 104, 113, 117, 123, 125, 130
<223> n = A,T,C or G

<400> 356
ctagtncagt gngntggaat tcgacaaaac accaaatggc ggatgacgcc ggtgcancgg 60
gggggcccgg gggccctggn ggccctgnga tggggaaccg cgnnggcttc cgnggangtt 120
tcngnagtgn catccggggc cggg 144

<210> 357
<211> 178
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 13
<223> n = A,T,C or G

<400> 357
ctagtcccct acngttaata tcactactaa ttaggctata accaggtctt tcctggcctg 60
agaaatattc tcttaaaatg acctttgttt taatctcatt catgatgttg attttttttc 120
aatgtggtgc aatatataca ataaaatttg tcataactat aaaaaaaaaa aaaagggc 178

<210> 358
<211> 471
<212> DNA
<213> Homo sapiens

<400> 358
ctagtaaaca acagcagcag aaacatcagt atcagcagcg tcgccagcag gagaatatgc 60
agcgccagag ccgaggagaa cccccgctcc ctgaggagga cctgtccaaa ctcttcaaac 120
caccacagcc gcctgccagg atggactcgc tgctcattgc aggccagata aacacttact 180
gccagaacat caaggagttc actgcccaaa acttaggcaa gctcttcatg gcccaggctc 240
ttcaagaata caacaactaa gaaaaggaag tttccagaaa agaagttaac atgaactctt 300
gaagtcacac cagggcaact cttggaagaa atatatttgc atattgaaaa gcacagagga 360
tttctttagt gtcattgccg attttggcta taacagtgtc tttctagcca taataaaata 420
aaacaaaatc ttgactgctt gctcatttga aaaaaaaaaa aaaaaaaggg c 471

<210> 359
<211> 285
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 130, 217, 251
<223> n = A,T,C or G

<400> 359
ctagtgacaa gctcctggtc ttgagatgtc ttctcgttaa ggagatgggc cttttggagg 60
taaaggataa aatgaatgag ttctgtcatg attcactatt ctagaacttg catgaccttt 120

actgtgttan ctctttgaat gttcttgaaa ttttaaactt tctttgtaaa caaatgatat 180
gtccttatca ttgtataaaa gctgttatgt gcaacantgt ggagattcct tgtctgattt 240
aataaaatac ntaaacactg aaaaaaaaaa aaaaaaaaaa agggc 285

<210> 360
<211> 280
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 125, 130, 144, 156, 179, 205, 206, 214
<223> n = A,T,C or G

<400> 360
ctagtgacaa gctcctggtc ttgagatgtc ttctcgttaa ggagatgggc cttttggagg 60
taaaggataa aatgaatgag ttctgtcatg attcactatt ctagaacttg catgaccttt 120
actgngttan ctctttgaat gttnttgaaa ttttanactt tctttgtaaa caaatgatnt 180
gtccttatca ttgtataaaa gctgnnatgt gcancagtgt ggagattcct tgtctgattt 240
aataaaatac ttaaacactg aaaaaaaaaa aaaaaagggc 280

<210> 361
<211> 374
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 351, 353
<223> n = A,T,C or G

<400> 361
ctagtgactt ttgtttagtg atagaagatt tggggaggac ccaaaggact cagaactttc 60
tctccatacc tccttttact cttttctttc tgtgtaatgt atcaacaact gtttaatctc 120
ccttctaaca aaccttgata taagctttct gatatcaaag tatattgaca gttaaccctt 180
actgatttta aacttgacta tccagtctgt taattaccta agattttgtt ttcatttcat 240
ctctaattgt tttgatcatt ggcagagaaa gagtatttga aattcatatc agttttgctc 300
cttattttaa tctctttgaa ttaaaaataa aacttttca aaatggaaaa nanaaaaaaa 360
aaaaaaaaaa gggc 374

<210> 362
<211> 199
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 195
<223> n = A,T,C or G

<400> 362
ctagtcacag ccctatactc cctctacata tttaccacaa cacaatgagg ctcactcacc 60
caccacatta acaacataaa accctcattc acacgagaaa acaccctcat gttcatacac 120
ctatccccca ttctcctcct atccctcaac cccgacatca ttaccgggtt ttcctcttaa 180
aaaaaaaaaa aaaangggg 199

<210> 363
<211> 500

<212> DNA
<213> Homo sapiens

<400> 363
ctagtctgca gatgtttctt gaatgctttg tcaaattaag aaagttaaag tgcaataatg 60
tttgaagaca ataagtggtg gtgtatcttg tttctaataa gataaacttt tttgtctttg 120
ctttatctta ttagggagtt gtatgtcagt gtataaaaca tactgtgtgg tataacaggc 180
ttaataaatt ctttaaaagg agagaactga aactagccct gtagatttgt ctggtgcatg 240
tgatgaaacc tgcagcttta tcggagtgat ggcaatcctc tgctggttta ttttcaagtg 300
gctgcgtttt ttttagtttg gcaggtgtag acttttaag ttgggcttta gaaaatctgg 360
gttagcctga agaaaattgc ctcagcctcc acagtaccat tttaaattca cataaaaggt 420
gaaagctcct ggttcagtgc catggcttca tggcattcag tgattagtgg taatggtaaa 480
cactggtgtg ttttgaagtt 500

<210> 364
<211> 206
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 40, 42, 57, 67, 68, 129, 162
<223> n = A,T,C or G

<400> 364
ctagttccag atctgaagcc caggttaggc atgacattgn ancccaacc ctacctnatc 60
tgtgctnnaa gacgctgaaa ctgcctggga tgttttcggg aacaagaatg tatatttgcc 120
ttatccctna acttggttta atcaaatcaa tgtgtgtatt anaataaaag tcacagcatc 180
aaaaaaaaaa aaaaaaaaaa aagggc 206

<210> 365
<211> 492
<212> DNA
<213> Homo sapiens

<400> 365
ctagtccagt gtggtggaat tcgaaccatg gagggtgtag aagagaagaa gaaggaggtt 60
cctgctgtgc cagaaaccct taagaaaaag cgaaggaatt tcgcagagct gaagatcaag 120
cgcctgagaa agaagtttgc ccaaaagatg cttcgaaagg caaggaggaa gcttatctat 180
gaaaaagcaa agcactatca caaggaatat aggcagatgt acagaactga aattcgaatg 240
gcgaggatgg caagaaaagc tggcaacttc tatgtacctg cagaacccaa attggcgttt 300
gtcatcagaa tcagaggtat caatggagtg agcccaaagg ttcgaaaggt gttgcagctt 360
cttcgccttc gtcaaatctt caatggaacc tttgtgaagc tcaacaaggc ttcgattaac 420
atgctgagga ttgtagagcc atatattgca tgggggtacc ccaatctgaa gtcagtaaat 480
gaactaatct ac 492

<210> 366
<211> 305
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 35, 38, 89, 202
<223> n = A,T,C or G

<400> 366
ctagtccagt gtggtggaat tccgtcctgc gcggntgntc tctggagcag cgttctttta 60

tctccgtccg ccttctctcc tacctaagng cgtgccgcca cccgatggaa gattcgatgg 120
acatggacat gagccccctg aggccccaga actatctttt cggttgtgaa ctaaaggccg 180
acaaagatta tcactttaag gnggataatg atgaaaatga gcaccagtta tctttaagaa 240
cggtcagttt aggggctggt gcaaaggatg agttgcacat tgttgaagca gaggcaatga 300
attac 305

<210> 367
<211> 508
<212> DNA
<213> Homo sapiens

<400> 367
ctagttttgt taggaacatt tgagttactt caatcatttt cacaggcagc caacaagcaa 60
ttaagagcag ttataataga ggaagctggg ggacccattt tgcaccatga gtttgtgaaa 120
aatctggatt aaaaaattac ctcttcagtg ttttctcatg caaaattttc ttctagcatg 180
tgataatgag taaactaaaa ctattttcag cttttctcaa ttaacatttt ggtagtatac 240
ttcagagtga tgttatctaa gtttaagtag tttaagtatg ttaaatgtgg atcttttaca 300
ccacatcaca gtgaacacac tggggagacg tgcttttttg gaaaactcaa aggtgctagc 360
tccctgattc aaagaaatat ttctcatgtt tgttcattct agtttatatt ttcatttaaa 420
atcctttagg ttaagtttaa gctttttaaa agttagtttt gagaattgag acacaatact 480
aatactgtag gaattggtga ggccttga 508

<210> 368
<211> 168
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 161
<223> n = A,T,C or G

<400> 368
ctagtgtgac aaaataacta catcctaatg aaaatcaagt ttgatatgtt tgttttgaaa 60
gtagcgttgg aagagttgtt gggggttttt tgcatccata gcactggtta ctttgaacaa 120
ataaataaaa gctttctgta gttgcttcct ttatcaaaaa naacattt 168

<210> 369
<211> 517
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 154
<223> n = A,T,C or G

<400> 369
ctagtatatg ggtaacaaat gaatatgtct gaacctcagc tataatactt tctactacct 60
ttgcaaggag atgggatagg aacaatcact cagaggaggc gttgcatggg cagggtcata 120
gggggaagaa aggtggttta gctgttttat ttanccattc agggggctct ccagagagga 180
gacggtggta gagggtgaac tagagaagat aagaatgtct tcctaggccg gatgcggtgg 240
ctcacgcctg taatcccagc actttgggat tgcgaggtgg gcggatcact tgaggtcagg 300
agttcaagac cagcctggcc aacatggtaa aacccgtctc tactaacaat acaaagatta 360
gcctggtgtg gtggcacggg cctgtaatcg cagcccctg gaaggccaag gcaggagaat 420
cgcctcaaca ctggaggtgg aggttgcagt gagctgaaat tgtgccactg cactccaccc 480
tgggcaatga ggcaagaccc tgtctcaaaa aataata 517

<210> 370
<211> 601
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 563
<223> n = A,T,C or G

<400> 370
ctagtgtgac aaaataacta catcctaatg aaaatcaagt ttgatatgtt tgttttgaaa 60
gtagcgttgg aagagttgtt gggggttttt tgcatccata gcactggtta ctttgaacaa 120
ataaataaaa gctttctgta gttgcttcct ttatcagaaa agaacatttg ataccatggt 180
atatcatttc ctcttcatta aagaacagct tttctaaatg ttgggggaaa tgtccatagt 240
cattactcag tcaaaacttg tgttctcatg agcctaagga ccattctaga tttattacgt 300
gtttttttt tgtgtgtgtg tgtgtgtgtg tgtgtatcca taaaatgcat atgtaaattt 360
ttttttgttt ttaagcattc acccaaacaa aaaaatcaca ggtaaaccca tgtttctgag 420
atgccattat tccaagcaaa ataagagata atcccttcaa gttaaattga aaattttcct 480
gaaaccatac atttcaagtg aaataagtaa ttctagatag gacaatttaa attggataat 540
tttaaagtgt ctataattgc agnggtttat ttgcaaaatt cctaaaagga aaaatttatc 600
a 601

<210> 371
<211> 555
<212> DNA
<213> Homo sapiens

<400> 371
ctagtgtgac aaaataacta catcctaatg aaaatcaagt ttgatatgtt tgttttgaaa 60
gtagcgttgg aagagttgtt gggggttttt tgcatccata gcactggtta ctttgaacaa 120
ataaataaaa gctttctgta gttgcttcct ttatcagaaa agaacatttg ataccatggt 180
atatcatttc ctcttcatta aagaacagct tttctaaatg ttgggggaaa tgtccatagt 240
cattactcag tcaaaacttg tgttctcatg agcctaagga ccattctaga tttattacgt 300
gtttttttt tgtgtgtgtg tgtgtgtgtg tgtgtatcca taaaatgcat atgtaaattt 360
ttttttgttt ttaagcattc acccaaacaa aaaaatcaca ggtaaaccca tgtttctgag 420
atgccattat tccaagcaaa ataagagata atcccttcaa gttaaattga aaattttcct 480
gaaaccatac atttcaagtg aaataagtaa ttctagatag gacaatttaa attggataat 540
tttaaagtgt ctata 555

<210> 372
<211> 418
<212> DNA
<213> Homo sapiens

<400> 372
ctagtttaag gagactggcc gaagctctgc ccaaacaatc tgtggatgga aaagcaccac 60
ttgctactgg agaggatgat gatgatgaag ttccagatct tgtggagaat tttgatgagg 120
cttccaagaa tgaggcaaac tgaattgagt caacttctga agataaaacc tgaagaagtt 180
actgggagct gctattttat attatgactg ctttttaaga aatttttgtt tatggatctg 240
ataaaatcta gatctctaat atttttaagc ccaagcccct tggacactgc agctcttttc 300
agtttttgct tatacacaat tcattctttg cagctaatta agccgaagaa gcctgggaat 360
caagtttgaa acaaagatta ataaagttct ttgcctagta aaaaaaaaaa aaaagggc 418

<210> 373
<211> 130
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 2, 12, 15, 16
<223> n = A,T,C or G

<400> 373
nngtgtgaca anctnnctac atcctaatga aaatcaagtt tgatatgttt gttttgaaag 60
tagcgttgga agagttgttg ggggtttttt gcatccatag cactggttac tttgaacaaa 120
taaataaaag 130

<210> 374
<211> 460
<212> DNA
<213> Homo sapiens

<400> 374
ctagtcctct tagaatttct tgcgctttga ttttttagg gcttgtgccc tgtttcactt 60
atagggtcta gaatgcttgt gttgagtaaa aaggagatgc ccaatattca aagctgctaa 120
atgttctctt tgccataaag actccgtgta actgtgtgaa cacttgggat ttttctcctc 180
tgtcccgagg tcgtcgtctg ctttcttttt tgggtttctt tctagaagat tgagaagtgc 240
atatgacagg ctgagagcac ctccccaaac acacaagctc tcagccacag gcagcttctc 300
cacagcccca gcttcgcaca ggctcctgga gggctgcctg ggggaggcag acatgggagt 360
gccaaggtgg ccagatggtt ccaggactac aatgtcttta tttttaactg tttgccactg 420
ctgccctcac ccctgcccgg ctctggagta ccgtctgccc 460

<210> 375
<211> 397
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 348, 371, 391
<223> n = A,T,C or G

<400> 375
ctagttttta tagctatcaa cattaggagt aactttcaac cttgccagca tcactggtat 60
gatgtatatt taattaaagc acacttttcc ccgaccgtat acttaaaatg acaaagccat 120
tcttttaaat atttgtgact ctttcctaaa gccaaagttt ctgttgaatt atgttttgac 180
acacccctaa gtacaaggtg gtatggttgt gtacacatgc tgccttcttg gggattcaaa 240
aacaggtttt tgattttgaa tagcaattag tgatatagtg ctgtttaagc tactaacgat 300
aaaaggtaat aacattttat acaatttcca tatagtctat tcattaanta atctttttac 360
agttgcatca ngcctgaacc cgtccattca naaagct 397

<210> 376
<211> 422
<212> DNA
<213> Homo sapiens

<400> 376
ctagttcagg ccttccagtt cactgacaaa catggggaag tgtgcccagc tggctggaaa 60
cctggcagtg ataccatcaa gcctgatgtc caaaagagca aagaatattt ctccaagcag 120
aagtgagcgc tgggctgttt tagtgccagg ctgcggtggg cagccatgag aacaaaacct 180
cttctgtatt tttttttttcc attagtaaaa cacaagactt cagattcagc cgaattgtgg 240
tgtcttacaa ggcaggcctt tcctacaggg ggtggagaga ccagcctttc ttcctttggt 300
aggaatggcc tgagttggcg ttgtgggcag gctactggtt tgtatgatgt attagtagag 360
caacccatta atcttttgta gtttgtatta aacttgaact gagaaaaaaa aaaaaaaagg 420

gc 422

<210> 377
<211> 198
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 163, 197
<223> n = A,T,C or G

<400> 377
ctagtatatt taaacttaca ggcttatttg taatgtaaac caccatttta atgtactgta 60
attaacatgg ttataatacg tacaatcctt ccctcatccc atcacacaac ttttttgtg 120
tgtgataaac tgattttggt ttgcaataaa accttgaaaa atntttaaaa aaaaaaaaaa 180
aaaaaaaaag ggggggnc 198

<210> 378
<211> 388
<212> DNA
<213> Homo sapiens

<400> 378
ctagtgcttc tggcgatgac atttctaagc tacagcgtac tccaggagaa gaaaagatta 60
ataccttaaa agaagaaaac actcaagaag cagcagtcct gaatggtgtt tcataaactg 120
aagaagttcc tagtttacag ttcttttaca ttacatttac aatagtgctt gtacaagctt 180
gccaaagata gaatatggat cgccagtctt tacatcgcac tttcagttcc tccatttgga 240
attcaaaaag gggagggatc ctgaagaaat catatgttaa acatactttg acacctactg 300
tgttataaaa tatatcatca gatgtgcctt gagaatagta tatgtaacat taaaaaaaag 360
ttgctggcta aaaaaaaaaa aaaagggc 388

<210> 379
<211> 277
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 254
<223> n = A,T,C or G

<400> 379
ctagttacaa aaataattta aggtgaaatc tctaatattt ataaaagtag caaaataaat 60
gcataattaa aatatatttg gacataacag acttggaagc agatgataca gacttctttt 120
tttcataatc aggttagtgt aagaaattgc catttgaaac aatccatttt gtaactgaac 180
cttatgaaat atatgtattt catggtacgt attctctagc acagtctgag caattaaata 240
gattcataag catnaaaaaa aaaaaaaaaa aaagggc 277

<210> 380
<211> 458
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 371
<223> n = A,T,C or G

<400> 380
ctagttatca gatcctttga aaagagaata tttacaatat atgactaatt tggggaaaat 60
gaagttttga tttatttgtg tttaaatgct gctgtcagac gattgttctt agacctccta 120
aatgccccat attaaaagaa ctcattcata ggaaggtgtt tcattttggt gtgcaaccct 180
gtcattacgt caacgcaacg tctaactgga cttcccaaga taaatggtac cagcgtcctc 240
ttaaaagatg ccttaatcca ttccttgagg acagacctta gttgaaatga tagcagaatg 300
tgcttctctc tggcagctgg ccttctgctt ctgagttgca cattaatcag attagcctgt 360
attctcttca ntgaattttg ataatggctt ccagactctt tggcgttgga gacgcctgtt 420
aggatcttca agtcccatca tagaaaattg aaacacaa 458

<210> 381
<211> 315
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 12
<223> n = A,T,C or G

<400> 381
ctagtccagt gnggtggaat tcgaggaatc agaaacctga agttagaaag gctcaacgag 60
aacaagctat cagggctgct aaggaagcaa aaaaggctaa gcaagcatct aaaaagactg 120
caatggctgc tgctaaggca cctacaaagg cagcacctaa gcaaaagatt gtgaagcctg 180
tgaaagtttc agctccccga gttggtggaa aacgctaaac tggcagatta gatttttata 240
atccaatctt tatttaaaaa tctaatctgc cagtttagat ttttaaataa agattggatt 300
ataaaaaaaa aaaaa 315

<210> 382
<211> 253
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 38, 158, 162
<223> n = A,T,C or G

<400> 382
ctagtgattt tgagtatgtt gttgattttt ttgtgtgngg ttactgatag aatcaagaca 60
attacaactt cataaatgac aaataatagg attatctcca cattttctgt tgctggagga 120
acaaaacatt gtgcccattt gaaaatttta atttttgntg gnttaactat cccacattat 180
aaatcatcct tcaccatttt atatcagtta aatatgggtg tgttggggag gaatgactgg 240
catgtagaca tgt 253

<210> 383
<211> 413
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 158, 199, 202, 207, 230, 273, 338, 351, 365
<223> n = A,T,C or G

<400> 383
ctagttttta tagctatcaa cattaggagt aactttcaac cttgccagca tcactggtat 60

gatgtatatt taattaaagc acacttttcc ccgaccgtat acttaaaatg acaaagccat 120
tcttttaaat atttgtgact ctttcctaaa gccaaagntt ctgttgaatt atgttttgac 180
acacccctaa gtacaaggng gnatggntgt gtacacatgc tgccttcttn gggggattca 240
aaaacaggtt tttgattttg aatagcaatt agngatatag tgctgtttaa gctactaacg 300
ataaaaggta ataacatttt atacaatttc catatagnct attcattaag naatcttttt 360
acagntgcat caggcctgaa cccgtccatt cagaaagctt caaattatag aaa 413

<210> 384
<211> 321
<212> DNA
<213> Homo sapiens

<400> 384
ctagtccagt gtggtggaat tcgaggaatc agaaacctga agttagaaag gctcaacgag 60
aacaagctat cagggctgct aaggaagcaa aaaaggctaa gcaagcatct aaaaagactg 120
caatggctgc tgctaaggca cctacaaagg cagcacctaa gcaaaagatt gtgaagcctg 180
tgaaagtttc agctccccga gttggtggaa aacgctaaac tggcagatta gatttttata 240
atccaatctt tatttaaaaa tctaatctgc cagtttagat ttttaaataa agattggatt 300
ataaaaaaaa aaaaaaaggg c 321

<210> 385
<211> 400
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 329, 376, 397
<223> n = A,T,C or G

<400> 385
ctagtgcttt acctttatta atgaactgtg acaggaagcc caaggcagtg ttcctcacca 60
ataacttcag agaagtcagt tggagaaaat gaagaaaaag gctggctgaa aatcactata 120
accatcagtt actggtttca gttgacaaaa tatataatgg tttactgctg tcattgtcca 180
tgcctacaga taatttattt tgtatttttg aataaaaaac atttgtacat tcctgatact 240
gggtacaaga gccatgtacc agtgtactgc tttcaactta aatcactgag gcatttttac 300
tactattctg ttaaaatcag gattttagng cttgccacca ccagatgaga aggtaagcag 360
cctttctgtg gagagngaga ataattgtgt acaaagnaga 400

<210> 386
<211> 524
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 453, 476, 493, 498
<223> n = A,T,C or G

<400> 386
ctagtccagt gtggtggaat tcgcttggag gttggcggcg cggggctgaa ggctagcaaa 60
ccgagcgatc atgtcgcaca aacaaattta ctattcggac aaatacgacg acgaggagtt 120
tgagtatcga catgtcatgc tgcccaagga catagccaag ctggtcccta aaacccatct 180
gatgtctgaa tctgaatgga ggaatcttgg cgttcagcag agtcagggat gggtccatta 240
tatgatccat gaaccagaac ctcacatctt gctgttccgg cgcccactac ccaagaaacc 300
aaagaaatga agctggcaag ctacttttca gcctcaagct ttacacagct gtccttactt 360
cctaacatct ttctgataac attattatgt tgccttcttg tttctcactt tgatatttaa 420
aagatgttca atacactgtt tgaatgtgct ggntaactgc tttgcttctt gagtanagcc 480

accaccacca tancccancc agatgagtgc tctgtggacc caca 524

<210> 387
<211> 279
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 275
<223> n = A,T,C or G

<400> 387
ctagtgacaa gctcctggtc ttgagatgtc ttctcgttaa ggagatgggc cttttggagg 60
taaaggataa aatgaatgag ttctgtcatg attcactatt ctagaacttg catgaccttt 120
actgtgttag ctctttgaat gttcttgaaa ttttagactt tctttgtaaa caaataatat 180
gtccttatca ttgtataaaa gctgttatgt gcaacagtgt ggagattcct tgtctgattt 240
aataaaatac ttaaacactg aaaaaaaaaa aaaangggg 279

<210> 388
<211> 463
<212> DNA
<213> Homo sapiens

<400> 388
ctagttttgt taggaacatt tgagttactt caatcatttt cacaggcagc caacaagcaa 60
ttaagagcag ttataataga ggaagctggg ggacccattt tgcaccatga gtttgtgaaa 120
aatctggatt aaaaaattac ctcttcagtg ttttctcatg caaaattttc ttctagcatg 180
tgataatgag taaactaaaa ctattttcag cttttctcaa ttaacatttt ggtagtatac 240
ttcagagtga tgttatctaa gtttaagtag tttaagtatg ttaaatgtgg atcttttaca 300
ccacatcaca gtgaacacac tggggagacg tgcttttttg gaaaactcaa aggtgctagc 360
tccctgattc aaagaaatat ttctcatgtt tgttcattct agtttatatt ttcatttaaa 420
atcctttagg ttaagtttaa gctttttaaa agttagtttt gag 463

<210> 389
<211> 402
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 341, 392
<223> n = A,T,C or G

<400> 389
ctagtcacta ctgtcttctc cttgtagcta atcaatcaat attcttccct tgcctgtggg 60
cagtggagag tgctgctggg tgtacgctgc acctgcccac tgagttgggg aaagaggata 120
atcagtgagc actgttctgc tcagagctcc tgatctaccc cacccctag gatccaggac 180
tgggtcaaag ctgcatgaaa ccaggccctg gcagcaacct gggaatggct ggaggtggga 240
gagaacctga cttctctttc cctctccctc ctccaacatt actggaactc tatcctgtta 300
ggatcttctg agcttgtttc cctgctgggt gggacagagg ncaaaggaga agggagggtc 360
tagaagaggc agcccttctt tgtcctctgg gnaaatgagc tt 402

<210> 390
<211> 374
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 126, 222, 224, 237
<223> n = A,T,C or G

<400> 390
ctagtcacta ctgtcttctc cttgtagcta atcaatcaat attcttccct tgcctgtggg 60
cagtggagag tgctgctggg tgtacgctgc acctgcccac tgagttgggg aaagaggata 120
atcagngagc actgttctgc tcagagctcc tgatctaccc caccccctag gatccaggac 180
tgggtcaaag ctgcatgaaa ccaggccctg gcagcaacct gngnaatggc tggaggnggg 240
agagaacctg acttctcttt ccctctccct cctccaacat tactggaact ctatcctgtt 300
aggatcttct gagcttgttt ccctgctggg tgggacagag gacaaaggag aagggagggt 360
ctagaagagg cagc 374

<210> 391
<211> 243
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 129, 136, 156, 165
<223> n = A,T,C or G

<400> 391
cggaacagga ctatcgtgcc ctgctgattg ctgatacgcc cattattgat gttcgcgccc 60
ctatcgagtt tgagcacggc gcaatgcccg ccgctatcaa tctgccgtta atgaataacg 120
atgaacgcnc cgccgntggc acctgctata aacagnaagg ctcanacgca gcgctggcgc 180
tgggacataa actggtggcg ggtgaaattc gtcagcagcg catggacgcc tggcgggcag 240
cgt 243

<210> 392
<211> 390
<212> DNA
<213> Homo sapiens

<400> 392
ctagtggtga atgcatgtgt ctgtctgatc agcatcactg cacacggagg tctagtgagc 60
ctcttgctaa gtgtcacaca cactcttccc aaagacgtga tgagttaaag ttgtattctg 120
aaatcatgaa gccagagcct gtgccagacc ttctgctacc tctcatagaa ttgctctgta 180
attctaaatt taaaattaga agtagagaga gataagccat cgccccttg cctctgagaa 240
ttggctgctg tttctaatat aattattttc taagatagcc agatagttag aaaaagattt 300
tcattgatga catatcttta aactttcttg catcagtatt ctaaattgag caaactgaaa 360
gattttcatc aggaaaggag cactgtggga 390

<210> 393
<211> 86
<212> DNA
<213> Homo sapiens

<400> 393
aggaacattt gagttacttc aatcattttc acaggcagcc aacaagcaat taagagcagt 60
tataatagag gaagctgggg gaccca 86

<210> 394
<211> 420
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 353, 376, 397, 405
<223> n = A,T,C or G

<400> 394
ctagtgcttt acctttatta atgaactgtg acaggaagcc caaggcagtg ttcctcacca 60
ataacttcag agaagtcagt tggagaaaat gaagaaaaag gctggctgaa aatcactata 120
accatcagtt actggtttca gttgacaaaa tatataatgg tttactgctg tcattgtcca 180
tgcctacaga taatttattt tgtatttttg aataaaaaac atttgtacat tcctgatact 240
gggtacaaga gccatgtacc agtgtactgc tttcaactta aatcactgag gcatttttac 300
tactattctg ttaaaatcag gattttagtg cttgccacca ccagatgaga agntaagcag 360
cctttctgtg gagagngaga ataattgtgt acaaagnaga gaagnatcca attatgtgac 420

<210> 395
<211> 283
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 156, 217
<223> n = A,T,C or G

<400> 395
ctagtgacaa gctcctggtc ttgagatgtc ttctcgttaa ggagatgggc cttttggagg 60
taaaggataa aatgaatgag ttctgtcatg attcactatt ctagaacttg catgaccttt 120
actgtgttag ctctttgaat gttcttgaaa ttttanactt tctttgtaaa caaataatat 180
gtccttatca ttgtataaaa gctgttatgt gcaacantgt ggagattcct tgtctgattt 240
aataaaatac ttaaacactg aaaaaaaaaa aaaaaaaaag ggc 283

<210> 396
<211> 213
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 14, 15, 118, 119, 188
<223> n = A,T,C or G

<400> 396
gagctctagg ctgnncaaat ttaaaaacta ctatgtgatt aactcgagcc tttagttttc 60
atccatgtac atggatcaca gtttgctttg atcttcttca atatgtgaat ttgggctnnc 120
agaatcaaag cctatgcttg gtttaatgct tgcaatctga gctcttgaac aaataaaatt 180
aactattngt agtgtgaaaa aaaaaaaaaa agg 213

<210> 397
<211> 66
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 2, 3, 42
<223> n = A,T,C or G

<400> 397
cnnctatagg gcgaattggg taccgggccc cccctcgagg tngacggtat cgataagctt 60
gatatc 66

<210> 398
<211> 288
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 225, 232, 241, 244
<223> n = A,T,C or G

<400> 398
gacaagctcc tggtcttgag atgtcttctc gttaaggaga tgggcctttt ggaggtaaag 60
gataaaatga atgagttctg tcatgattca ctattctaga acttgcatga cctttactgt 120
gttagctctt tgaatgttct tgaaatttta gactttcttt gtaaacaaat gatatgtcct 180
tatcattgta taaaagctgt tatgtgcaaa aaaaaaaaaa aaaangggcg gncgccaccg 240
nggntggagc tccagctttt gttcccttta gtgagggtta attgccgc 288

<210> 399
<211> 156
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 107, 108
<223> n = A,T,C or G

<400> 399
aaatttaaaa actactatgt gattaactcg agcctttagt tttcatccat gtacatggat 60
cacagtttgc tttgatcttc ttcaatatgt gaatttgggc tcacagnntc aaagcctatg 120
cttggtttaa tgcttgcaat ctgagctctt gaacaa 156

<210> 400
<211> 551
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 83, 221, 237, 338, 350, 359, 519, 542
<223> n = A,T,C or G

<400> 400
tggaattctg catctgtatc cagcgccagg tcccgccagt cccagctgcg cgcgcccccc 60
agtcccgcac ccgttcggcc cangctaagt tagccctcac catgccggtc aaaggaggca 120
ccaagtgcat caaatacctg ctgttcggat ttaacttcat cttctggctt gccgggattg 180
ctgtccttgc cattggacta tggctccgat tcgactctca naccaagagc atcttcnagc 240
aagaaactaa taataataat tccagcttct acacaggagt ctatattctg atcggagccg 300
gcgccctcat gatgctggtg ggcttcctgg gctgctgngg ggctgtgcan gagtcccant 360
gcatgctggg actgttcttc ggcttcctct tggtgatatt cgccattgaa atagctgcgg 420
ccatctgggg atattcccac aaggatgagg tgattaagga agtccaggag tttttacaag 480
gacacctaca acaagctgaa aaccaaggat gagccccanc ggggaaacgc tgaaaagcca 540
tncactatgc g 551

<210> 401
<211> 157
<212> DNA
<213> Homo sapiens

<400> 401
aggatagaaa cactgtgtcc cgagagtaag gagagaagct actattgatt agagcctaac 60
ccaggttaac tgcaagaaga ggcgggatac tttcagcttt ccatgtaact gtatgcataa 120
agccaatgta gtccagtttc taagatcatg ttccaag 157

<210> 402
<211> 546
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 534
<223> n = A,T,C or G

<400> 402
gtaactcctt catgcaataa actgaaaaga gccatgctgt ctagtcttga agtccctcat 60
ttaaacagag gtcaagcaat aggcgcctgg cagtgtcaag cctgaaacca agcaataccg 120
tcatgtttca gccaagccca gagccctaag attacaaaca actatggccg gaacctcctc 180
agctctccct ctgcagagtt ccctacccta agagaatgtt accacctgaa cagtcctcgg 240
tgaatctgag aggagaggat ggggtaaggc agaagcacca gctgtactac tagaagggag 300
cttttggtgg tagatccccc ggtgtctcca acctgactag gtggacagag ctcaaagagg 360
ccctcttacc gctagcgagg tgataggaca tctggcttgc cacaaaggtc tgttcgacca 420
gacatatcct agctaaggga tgtccaaaca tcagaatgtt gaggccaacc ttcctatcag 480
agttaaactt tttgacaagg gaacaaatct caaactgatc catcagtcat gtanctagct 540
gtagag 546

<210> 403
<211> 579
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 305, 523, 532
<223> n = A,T,C or G

<400> 403
tttgcaaata ttcccctggt agcctacttc cttacccccg aatattggta agatcgagca 60
atggcttcag gacatgggtt ctcttctcct gtgatcattc aagtgctcac tgcatgaaga 120
ctggcttgtc tcagtgtttc aacctcacca gggctgtctc ttggtccaca cctcgctccc 180
tgttagtgcc gtatgacagc ccccatcaaa tgaccttggc caagtcacgg tttctctgtg 240
gtcaaggttg gttggctgat tggtggaaag tagggtggac caaaggaggc cacgtgagca 300
gtcancacca gttctgcacc agcagcgcct ccgtcctagt gggtgttcct gtttctcctg 360
gccctgggtg ggctagggcc tgattcggga agatgccttt gcagggaggg gaggataagt 420
gggatctacc aattgattct ggcaaaacaa tttctaagat ttttttgctt ttatgtggga 480
aacagatcta aaatctcatt ttatgctgta ttttatatct tanttgtgtt tngaaaacgt 540
ttttgatttt tggaaacaca tcaaaataaa taatggcgt 579

<210> 404
<211> 599
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<222> 32, 33
<223> n = A,T,C or G

<400> 404
tggaattcga acgtatggtc caggaagctg annagtacaa agctgaagat gagaagcaga 60
gggacaaggt gtcatccaag aattcacttg agtcctatgc cttcaacatg aaagcaactg 120
ttgaagatga gaaacttcaa ggcaagatta acgatgagga caaacagaag attctggaca 180
agtgtaatga aattatcaac tggcttgata agaatcagac tgctgagaag gaagaatttg 240
aacatcaaca gaaagagctg gagaaagttt gcaaccccat catcaccaag ctgtaccaga 300
gtgcaggagg catgccagga ggaatgcctg ggggatttcc tggtggtgga gctcctccct 360
ctggtggtgc ttcctcaggg cccaccattg aagaggttga ttaagccaac caagtgtaga 420
tgtagcattg ttccacacat ttaaaacatt tgaaggacct aaattcgtag caaattctgt 480
ggcagttttt aaaaagttta agctgctata gtaaagttta ctgggcattc tcaatacttg 540
aatatggaac atatgcacag ggggaaggaa taacattgca ctttataaac actgtattg 599

<210> 405
<211> 204
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 51, 76, 77, 91, 92, 98
<223> n = A,T,C or G

<400> 405
aaataatacg aaactttaaa aagcattgga gtgtcagtat gttgaatcag nagtttcact 60
ttaactgtaa acaatnnctt aggacaccat nngggctngt ttctgtgtaa gtgtaaatac 120
tacaaaaact tatttatact gttcttatgt catttgttat attcatagat ttatatgatg 180
atatgacatc tggctaaaaa agaa 204

<210> 406
<211> 414
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 149, 263, 271, 304, 390
<223> n = A,T,C or G

<400> 406
aatgcatcaa cataatttct gtattaacca tcatgcgcac aagaaataca tagtaaataa 60
ggaagctgaa aactcctggc attggatctt aagctagatg attagaatgt gaaaaagatt 120
ttacaaatgt aaaacttcta tttctctgna gaaactttct tcactttgct gtgcaagaag 180
acactgcttt gctatattta aaatggcttt tttaaaagag atttatgtat ttggtaaatg 240
tttgtagtca acagttcaca cangaagctg ntacacggtt tgatcatgta aaaccgtttt 300
ggcnggcaca agctggactt tgttgccatc cttgagatga acctttaag aaaaataagt 360
taatctcaat ttttccctga atgtgtttgn ttttcttcat tatacaataa atat 414

<210> 407
<211> 412
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 132, 264, 272, 358, 386, 390
<223> n = A,T,C or G

<400> 407
naatgcatca acataatttc tgtattaacc atcatgcgca caagaaatac atagtaaata 60
aggaagctga aaactcctgg cattggatct taagctagat gattagaatg tgaaaaagat 120
tttacaaatg tnaaacttct atttctctgt agaaactttc ttcactttgc tgtgcaagaa 180
gacactgctt tgctatattt aaaatggctt ttttaaaaga gatttatgta tttggtaaat 240
gtttgtagtc aacagttcac acangaagct gnacacggtt tgatcatgta aaaccgtttg 300
gcggcacaag ctggactttg ttgccatcct tgagatgaac cttttaagaa aaataagnta 360
atctcaattt tttccctgaa tgtgtngttn ttcttcatta tacaataaat at 412

<210> 408
<211> 568
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 446, 478, 500, 502, 514, 533, 543
<223> n = A,T,C or G

<400> 408
tttagccaag gctgtggcaa aggtgtaact tgtaaacttg agttggagta ctatatttac 60
aaataaaatt ggcaccatgt gccatctgta catattactg ttgcatttac ttttaataaa 120
gcttgtggcc ccttttactt ttttatagct taactaattt gaatgtggtt acttcctact 180
gtagggtagc ggaaaagttg tcttaaaagg tatggtgggg atatttttaa aaactccttt 240
tggtttacct ggggatccaa ttgatgtata tgtttatata ctgggttctt gttttatata 300
cctggctttt actttattaa tatgagttac tgaaggtgat ggaggtattt gaaaatttta 360
cttccatagg acatactgca tgtaagccaa gtcatggaga atctgctgca tagctctatt 420
ttaaagtaaa agtctaccac cgaatnccta ggtccccctg ttttctgttt cttcttgnga 480
ttgctgccat aatttctaan tnatttactt ttancactat ttaagttatc aantttagct 540
agnatcttca aactttcact ttgaaaaa 568

<210> 409
<211> 401
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 10, 102, 103, 376
<223> n = A,T,C or G

<400> 409
aaataatacn aaactttaaa aagcattgga gtgtcagtat gttgaatcag tagtttcact 60
ttaactgtaa acaatttctt aggacaccat ttgggctagt tnntgtgtaa gtgtaaatac 120
tacaaaaact tatttatact gttcttatgt catttgttat attcatagat ttatatgatg 180
atatgacatc tggctaaaaa gaaattattg caaaactaac cactatgtac ttttttataa 240
atactgtatg gacaaaaaat ggcatttttt atattaaatt gtttagctct ggcaaaaaaa 300
aaaaatttta agagctggta ctaataaagg attattatga ctgttaaaaa aaaaaaaaaa 360
gggcggccgc caccgnggtg gagctccagc ttttgttccc t 401

<210> 410
<211> 576

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 268, 386, 387, 421, 445, 447, 449, 456, 469, 500, 502, 541, 549, 569
<223> n = A,T,C or G

<400> 410
tggaattccg cttgccagcg tgttggagag accgctaccg gtgaaccagc gcgggttttt 60
cggacttggg ggtcgtgcag atctgctgga tctaggtcca gggagtctca gtgatggtct 120
gagcctggcc gcgccaggct ggggtgtccc agaagagcca ggaatcgaaa tgcttcatgg 180
aacaaccacc ctggccttca agttccgcca tggagtcata gttgcagctg actccagggc 240
tacagcgggt gcttacattg cctcccanac ggtgaagaag gtgatagaga tcaacccata 300
cctgctaggc accatggctg ggggcgcagc ggattgcagc ttctgggaac ggctgttggc 360
tcggcaatgt cgaatctatg agcttnnaaa taaggaacgc atctctgtag caagctgcct 420
ncaaactgct tgccaacatg gtgtntnant acaaangcat ggggctgtnc atgggcacca 480
tgatctgtgg ctgggataan anaggccctg gcctctacta cgtggacagt gaagggaacc 540
ngatttcang ggccaccttc tctgtaagnt ctggct 576

<210> 411
<211> 557
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1
<223> n = A,T,C or G

<400> 411
nccaacacag tcagaaacat tgttttgaat cctctgtaaa ccaaggcatt aatcttaata 60
aaccaggatc catttaggta ccacttgata taaaaaggat atccataatg aatattttat 120
actgcatcct ttacattagc cactaaatac gttattgctt gatgaagacc tttcacagaa 180
tcctatggat tgcagcattt cacttggcta cttcataccc atgccttaaa gaggggcagt 240
ttctcaaaag cagaaacatg ccgccagttc tcaagttttc ctcctaactc catttgaatg 300
taagggcagc tggccccccaa tgtggggagg tccgaacatt ttctgaattc ccattttctt 360
gttcgcggct aaatgacagt ttctgtcatt acttagattt ccgatctttc ccaaaggtgt 420
tgatttacaa agaggccagc taatagcaga aatcatgacc ctgaaagaga gatgaaattc 480
aagctgtgag ccaggcagga gctcagttat ggcaaaaggt tctttgagaa tcagccattt 540
ggtacaaaaa agatttt 557

<210> 412
<211> 499
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 455, 482
<223> n = A,T,C or G

<400> 412
gtaactcctt catgcaataa actgaaaaga gccatgctgt ctagtcttga agtccctcat 60
ttaaacagag gtcaagcaat aggcgcctgg cagtgtcaag cctgaaacca agcaataccg 120
tcatgtttca gccaagccca gagccctaag attacaaaca actatggccg gaacctcctc 180
agctctccct ctgcagagtt ccctacccta agagaatgtt accacctgaa cagtcctcgg 240

tgaatctgag aggagaggat ggggtaaggc agaagcacca gctgttacta ctagaaggga 300
gcttttggtg gtagatcccc tggtgtctcc aacctgacta ggtggacaga gctcaaagag 360
gccctcttac cgctagcgag gtgataggac atctggcttg ccacaaaggt tctgtttcga 420
ccagacatat cctagctaag ggatgtccaa acatnagaat gtgaggccaa accttctatc 480
anagttaaac ttttgacaa 499

<210> 413
<211> 238
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 100, 129, 130, 131, 159
<223> n = A,T,C or G

<400> 413
ggatagaaac actgtgtccc gagagtaagg agagaagcta ctattgatta gagcctaacc 60
caggttaact gcaagaagag gcgggatact ttcagctttn catgtaactg tatgcataaa 120
gccaatgtnn nccagtttct aagatcatgt tccaagctna ctgaatccca cttcaataca 180
cactcatgaa ctcctgatgg aacaataaca ggcccaagcc tgtggtatga tgtgcaca 238

<210> 414
<211> 279
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 169, 170, 183, 187, 235
<223> n = A,T,C or G

<400> 414
atatgggtaa caaatgaata tgtctgaacc tcagctataa tactttctac taccttttgca 60
aggagatggg ataggaacaa tcactcagag gaggcgttgc atgggcaggg tcatagggg 120
aagaaaggtg gtttagctgt tttatttagc cattcagggg gctctccann gaggagacag 180
gtngtanagg gtgaactagg agaagataag aatgtcttcc taggccggat gcggnggctc 240
acgcctgtaa tcccagcact ttgggattgc gaggtgggc 279

<210> 415
<211> 574
<212> DNA
<213> Homo sapiens

<400> 415
ccaacacagt cagaaacatt gttttgaatc ctctgtaaac caaggcatta atcttaataa 60
accaggatcc atttaggtac cacttgatat aaaaaggata tccataatga atattttata 120
ctgcatcctt tacattagcc actaaatacg ttattgcttg atgaagacct ttcacagaat 180
cctatggatt gcagcatttc acttggctac ttcataccca tgccttaaag aggggcagtt 240
tctcaaaagc agaaacatgc cgccagttct caagttttcc tcctaactcc atttgaatgt 300
aagggcagct ggcccccaat gtggggaggt ccgaacattt tctgaattcc catttcttg 360
ttcgcggcta aatgacagtt tctgtcatta cttagattcc gatctttccc aaaggtgttg 420
atttacaaag aggccagcta atagcagaaa tcatgaccct gaaagagaga tgaaattcaa 480
gctgtgagcc aggcaggagc tcagtatggc aaaggttctt gagaatcagc catttggtac 540
aaaaaagatt tttaaagctt ttatgttata ccat 574

<210> 416
<211> 545

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 533
<223> n = A,T,C or G

<400> 416
tggaattcct ttaaaccctg cgtggcaatc cctgacgcac cgccgtgatg cccagggaag 60
acagggcgac ctggaagtcc aactacttcc ttaagatcat ccaactattg gatgattatc 120
cgaaatgttt cattgtggga gcagacaatg tgggctccaa gcagatgcag cagatccgca 180
tgtcccttcg cgggaaggct gtggtgctga tgggcaagaa caccatgatg cgcaaggcca 240
tccgagggca cctggaaaac aacccagctc tggagaaact gctgcctcat atccgggga 300
atgtgggctt tgtgttcacc aaggaggacc tcactgagat cagggacatg ttgctggcca 360
ataaggtgcc agctgctgcc cgtgctggtg ccattgcccc atgtgaagtc actgtgccag 420
cccagaacac tggtctcggg cccgagaaga cctccttttt ccaggcttta ggtatcacca 480
ctaaaatctc caggggcacc attgaaatcc tgagtgatgt gcagctgatc aanactggag 540
acaaa 545

<210> 417
<211> 373
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 16, 17, 360, 361
<223> n = A,T,C or G

<400> 417
natttttta gattanntgt ctttaggtga tttaatggta ctttaataac tactaagaaa 60
tattggctat ttcaatgtaa gttataaggt ggtacattcc taagggtatt tatagttgat 120
gataacatga aaactgaaat aagataaaat acaacgtgct aaatctttta tgtattctaa 180
ctttaaaaga caagtgcaac aaagttagac tgacttctat atgtgctctt ttactctgat 240
aatattaaat taggactaac ttatgtttta taatgattat aatttacatg cttattttta 300
aaatagtata tgtggacaca tatatatcat tatattaaaa taaattctac cattttaaan 360
naaaagaaaa aaa 373

<210> 418
<211> 291
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 22, 23, 213, 217
<223> n = A,T,C or G

<400> 418
naggatagaa acactgtgtc cnnagagtaa ggagagaagc tactattgat tagagcctaa 60
cccaggttaa ctgcaagaag aggcgggata ctttcagctt tccatgtaac tgtatgcata 120
aagccaatgt agtccagttt ctaagatcat gttccaagct aactgaatcc cacttcaata 180
cacactcatg aactcctgat ggaacaataa canggcncca agcctgtggt atgatgtgca 240
cacttgctag actcagaaaa aatactactc tcataaatgg gtgggagtat t 291

<210> 419
<211> 596

<212> DNA
<213> Homo sapiens

<400> 419
agcctgcttt ggcagtgtgg ctttttgcac acttgccctg tcttcctgag actacttcag 60
taagccatgc ttccttcttc cccactttta tttggtgtca tgaatagaaa cttccaaatg 120
taaccatgga agctaagttt ggcctgcttt gcttttagt ctccacacca tgggcagaac 180
tgctgtcttt actacttcat ctcacccaag tcccgttccc aggcagccag gggcctgggt 240
ttgaataatt gcagggccag cctgccatga tctttctcac ttactcctct cccattcagc 300
aatcaaccag actaaggagt tttgatccct agtgattaca gccctgaaga aaattaaatc 360
tgaattaatt ttacatggcc ttcgtgatct ttctgctgtt cttacttttt cgaatgtagt 420
tgggggtgg gagggacagg ttatggtatt taaagagaat aaacattttg cacatacatg 480
tattgtacaa cagtaagatc ctctgttaaa accagctgtc ctgttctcca tctccatttc 540
ttcccatgct gtaaccccag gctccaccag ctgttcccca gtgatgttac ctagct 596

<210> 420
<211> 415
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 2, 3, 404, 405
<223> n = A,T,C or G

<400> 420
nnntggaatt cgcaagatgg cgggtgaaaa agttgagaag ccagatacta aagagaagaa 60
acccgaagcc aagaaggttg atgctggtgg caaggtgaaa aagggtaacc tcaaagctaa 120
aaagcccaag aaggggaagc cccattgcag ccgcaaccct gtccttgtca gaggaattgg 180
caggtattcc cgatctgcca tgtattccag aaaggccatg tacaagagga agtactcagc 240
cgctaaatcc aaggttgaaa agaaaaagaa ggagaaggtt ctcgcaactg ttacaaaacc 300
agttggtggt gacaagaacg gcggtacccg ggtggttaaa cttcgcaaaa tgcctagata 360
ttatcctact gaagatgtgc ctcgaaagct gttgagccac gggnnaaaaa ccctt 415

<210> 421
<211> 572
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 323, 524
<223> n = A,T,C or G

<400> 421
tggaattcct ttaaaccctg cgtggcaatc cctgacgcac cgccgtgatg cccagggaag 60
acagggcgac ctggaagtcc aactacttcc ttaagatcat ccaactattg gatgattatc 120
cgaaatgttt cattgtggga gcagacaatg tgggctccaa gcagatgcag cagatccgca 180
tgtcccttcg cgggaaggct gtggtgctga tgggcaagaa caccatgatg cgcaaggcca 240
tccgagggca cctggaaaac aacccagctc tggagaaact gctgcctcat atccgggga 300
atgtgggctt tgtgttcacc aangaggacc tcactgagat cagggacatg ttgctggcca 360
ataaggtgcc agctgctgcc cgtgctggtg ccattgcccc atgtgaagtc actgtgccag 420
cccagaacac tggtctcggg cccgagaaga cctccttttt ccaggcttta ggtatcacca 480
ctaaaatctc caggggcacc attgaaatcc tgagtgatgt gcanctgatc aagactggag 540
acaaagtggg agccagcgaa gccacgctgc tg 572

<210> 422
<211> 535

<212> DNA
<213> Homo sapiens

<400> 422
ccagtgtggt ggaattcaca gaagccacct tttttcattc tttcatttta aaaaaaagtg 60
agatatccac attccataaa attcaccctt tgaaagtaca caatgcaagt ttttaatata 120
ttcacaagtt tgtttaatcc ttaccactgt ctaattcaag agtattatca ttaccccaaa 180
aagaaaccca ttagcagtca ctccgcattc tcaccttccc ccatttcctc ccaaccacta 240
agtgattttc tgtctctatg gatttgcata ttctggacat tttatagaaa tggaatcatg 300
caatatatga tcttttgtgt ctggtgtctt tcaatgaaca atattgtcag tcttcatcca 360
cactgaagct tgtatcagta gtgagtgctt ccttttttatg gcggcatact aatccattgg 420
atggctatcc gacatttgtt ttatctatgc atcaattgca gtgagcctgg aggtggaaga 480
ctctggtttt tttagtgagc ccttcaagaa ggtacacatc ctggtgagag gatga 535

<210> 423
<211> 435
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 37, 39, 155, 243, 351, 367
<223> n = A,T,C or G

<400> 423
ccagtgtggt ggaattcctc gtctcaggcc agttgcngnc ttctcagcca aacgccgacc 60
aaggaaaact cactaccatg agaattgcag tgatttgctt ttgcctccta ggcatcacct 120
gtgccatacc agttaaacag gctgattctg gaagntctga ggaaaagcag ctttacaaca 180
aatacccaga tgctggggcc acatggctaa accctgaccc atctcagaag cagaatctcc 240
tanccccaca gaatgctgtg tcctctgaag aaaccaatga ctttaaacaa gagacccttc 300
caagtaagtc caacgaaagc catgaccaca tggatgatat ggatgatgaa natgatgatg 360
accatgngga caggcaggac tccattgact cgaacgactc tgatgatgta gatgacactg 420
atgattctca ccagt 435

<210> 424
<211> 558
<212> DNA
<213> Homo sapiens

<400> 424
ccagtgtggt ggaattcgca tcttctgagg tcaattaaaa ggagaaaaaa tacaatttct 60
cactttgcat ttagtcaaaa gaaaaaatgc tttatagcaa aatgaaagag aacatgaaat 120
gcttctttct cagtttattg gttgaatgtg tatctatttg agtctggaaa taactaatgt 180
gtttgataat tagtttagtt tgtggcttca tggaaactcc ctgtaaacta aaagcttcag 240
ggttatgtct atgttcattc tatagaagaa atgcaaacta tcactgtatt ttaatatttg 300
ttattctctc atgaatagaa atttatgtag aagcaaacaa aatactttta cccacttaaa 360
aagagaatat aacattttat gtcactataa tcttttgttt tttaagttag tgtatatttt 420
gttgtgatta tctttttgtg gtgtgaataa atcttttatc ttgaatgtaa taagaatttg 480
gtggtgtcaa ttgcttattt gttttcccac ggttgtccag caattaataa aacataacct 540
tttttactgc ctaaaaaa 558

<210> 425
<211> 600
<212> DNA
<213> Homo sapiens

<400> 425
tcatagccca tatatggagt tccgcgttac ataacttacg gtaaatggcc cgcctggctg 60

accgcccaac gacccccgcc cattgacgtc aataatgacg tatgttccca tagtaacgcc 120
aatagggact ttccattgac gtcaatgggt ggagtattta cggtaaactg cccacttggc 180
agtacatcaa gtgtatcata tgccaagtac gccccctatt gacgtcaatg acggtaaatg 240
gcccgcctgg cattatgccc agtacatgac cttatgggac tttcctactt ggcagtacat 300
ctacgtatta gtcatcgcta ttaccatggt gatgcggttt tggcagtaca tcaatgggcg 360
tggatagcgg tttgactcac ggggatttcc aagtctccac cccattgacg tcaatgggag 420
tttgttttgg caccaaaatc aacgggactt tccaaaatgt cgtaacaact ccgccccatt 480
gacgcaaatg ggcggtaggc gtgtacggtg ggaggtctat ataagcagag ctctctggct 540
aactagagaa cccactgctt actggcttat cgaaattaat acgactcact atagggagac 600

<210> 426
<211> 467
<212> DNA
<213> Homo sapiens

<400> 426
ccagtgtggt ggaattcaat aactaaaagg tatgcaatca aatctgcttt ttaaagaatg 60
ctctttactt catggacttc cactgccatc ctcccaaggg gcccaaattc tttcagtggc 120
tacctacata caattccaaa cacatacagg aaggtagaaa tatctgaaaa tgtatgtgta 180
agtattctta tttaatgaaa gactgtacaa agtagaagtc ttagatgtat atatttccta 240
tattgttttc agtgtacatg gaataacatg taattaagta ctatgtatca atgagtaaca 300
ggaaaatttt aaaaatacag atagatatat gctctgcatg ttacataaga taaatgtgct 360
gaatggtttt caaaataaaa atgaggtact ctcctggaaa tattaagaaa gactatctaa 420
atgttgaaag accaaaaggt taataaagta attataacta aaaaaaa 467

<210> 427
<211> 211
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 2, 9, 23, 30, 47, 72, 137
<223> n = A,T,C or G

<400> 427
gngcccacnc aggcaagctt tanagaaagn ggttgctgaa aataaanaaa tccagaaatt 60
ggcagagcag tntgtcctcc tcaatctggt ttatgaaaca actgacaaac acctttctcc 120
tgatggccat gtatgtnccc aggattatgt ttgttgaccc atctctgaca gttagagccg 180
atatcactgg aagatattca aaccgtctct a 211

<210> 428
<211> 615
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 496
<223> n = A,T,C or G

<400> 428
gggtactcaa cactgagcag atctgttctt tgagctaaaa accatgtgct gtaccaagag 60
tttgctcctg gctgctttga tgtcagtgct gctactccac ctctgcggcg aatcagaagc 120
aagcaacttt gactgctgtc ttggatacac agaccgtatt cttcatccta aatttattgt 180
gggcttcaca cggcagctgg ccaatgaagg ctgtgacatc aatgctatca tctttcacac 240
aaagaaaaag ttgtctgtgt gcgcaaatcc aaaacagact tgggtgaaat atattgtgcg 300

tctcctcagt aaaaaagtca agaacatgta aaaactgtgg cttttctgga atggaattgg 360
acatagccca agaacagaaa gaaccttgct ggggttggag gtttcacttg cacatcatgg 420
agggtttagt gcttatctaa tttgtgcctc actggacttg tccaattaat gaagttgatt 480
catattgcat catagnttgc tttgtttaag catcacatta aagttaaact gtattttatg 540
ttatttatag ctgtaggttt tctgtgttta gctatttaat actaattttc cataagctat 600
tttggtttag tgcaa 615

<210> 429
<211> 274
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 168
<223> n = A,T,C or G

<400> 429
ttttaagatc agagttcact ttctttggac tctgcctata ttttcttacc tgaacttttg 60
caagttttca ggtaaacctc agctcaggac tgctatttag ctcctcttaa gaagattaaa 120
agagaaaaaa aaaggccctt ttaaaaatag tatacactta ttttaagnga aaagcagaga 180
attttattta tagctaattt tagctatctg taaccaagat ggatgcaaag aggctagtgc 240
ctcagagaga actgtacggg gtttgtgact ggaa 274

<210> 430
<211> 690
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 11, 662
<223> n = A,T,C or G

<400> 430
ccagtgtggt ngaattcatc caggggggcta cccctggctc tctgttgcca gtggtcatca 60
tcgcagtggg tgtcttcctc ttcctggtgg cttttgtggg ctgctgcggg gcctgcaagg 120
agaactattg tcttatgatc acgtttgcca tctttctgtc tcttatcatg ttggtggagg 180
tggccgcagc cattgctggc tatgtgttta gagataaggt gatgtcagag tttaataaca 240
acttccggca gcagatggag aattacccga aaaacaacca cactgcttcg atcctggaca 300
ggatgcaggc agattttaag tgctgtgggg ctgctaacta cacagattgg gagaaaatcc 360
cttccatgtc gaagaaccga gtccccgact cctgctgcat taatgttact gtgggctgtg 420
ggattaattt caacgagaag gcgatccata aggagggctg tgtggagaag attgggggct 480
ggctgaggaa aaatgtgctg gtggtagctg cagcagccct tggaattgct tttgtcgagg 540
ttttgggaat tgtctttgcc tgctgcctcg tgaagagtat cagaagtggc tacgaggtga 600
tgtaaggggt ctggtctcct cagcctcctc atctgggggg agtggaatag tatcctccag 660
gnttttttcaa ttaaacggat tatttttca 690

<210> 431
<211> 155
<212> DNA
<213> Homo sapiens

<400> 431
tgcgggccgt attagaagca gtggggtacg ttagactcag atggaaaagt attctaggtg 60
ccagtgttag gatgtcagtt ttacaaaata atgaagcaat tagctatgtg attgagagtt 120
attgtttggg gatgtgtgtt gtggttttgc ttttt 155

<210> 432
<211> 233
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 1, 18
<223> n = A,T,C or G

<400> 432
nagtacataa ctacatantg ccaactctgg aatcaaattt ccttgtttga atcctgggac 60
cctattgcat taaagtacaa atactatgta tttttaatct atgatggttt atgtgaatag 120
gattttctca gttgtcagcc atgacttatg tttattacta aataaacttc aaactcctgt 180
tgaacattgt gtataactta gaataatgaa atataaggag tatgtgtaga aaa 233

<210> 433
<211> 271
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 182, 226
<223> n = A,T,C or G

<400> 433
ctgaccctgg gatctcctgt gctagcggcc aatgacaaat ccagtcattg gccaccagcc 60
acctctgcag tggggaccac actagcagcc ctgactccac actcctcctg gggacccaag 120
aggcagtgtt gctgactgcg tgtccacctt ggaatctggc tgaactggct gggaggacca 180
anactgcggc tggggtgggc agggaaggga agccgggggc tgctgngagg gatcttggag 240
cttccctgta gcccaccttc cccttgcttc a 271

<210> 434
<211> 438
<212> DNA
<213> Homo sapiens

<400> 434
aattccactc ctcccttgat ctttttggtt gtactttaat taagccctgc gagaatgctg 60
gataaatgcc ttgaagttag cagggtgtat ttttttagcg aatatgattt gcatgtcttg 120
ccaggagtta agcggcctct ggggtgttgg ggaaatactt tatttctttc catttatttt 180
ttgtggggcg gggatagggg agggcattga agttctacaa ttctggaata gttagttgat 240
ggtacatagt taacttggct tcggttacat attggacttt aacaactgaa gaatctatgc 300
gtgtcattta aagaaaagtt gcagaacaag caattggctt agatatacaa tctggaaaaa 360
tattcctgtg cccatatttt aatgtaattg tataactggg agcaaaaata tattctgctt 420
ttcaactgta ggtgctcc 438

<210> 435
<211> 500
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 203, 484
<223> n = A,T,C or G

<400> 435
catcgatggc atttcagtct ataggtaaac ttcctggaag ctggatttgg agacagttta 60
tcatctgatt attgggcttt cgtataggtc cttagggagc agcttacctg aaatgcattt 120
agtgtacacc agtctgtaaa cttcaacctg taatgaaagt gtaataaatg tacattgagt 180
tgatgtgata atgtgatata atnagaaata tatatttgat cttcctatct agttccttgt 240
tcagagctcc taaaaccctt gtaatttcca aagtgatgga gtacatcttt tgttctagta 300
tttggtcttt gaccccagtt cctgacacaa agctcctaaa ttcctttaaa tttcccagtg 360
ataggagaat tttttgttct aatgaggtca ctcttgatgg gcacctggat aactcaggat 420
gggggctgct cacaaagacc acatcatgat tggaagtttc aaactttcag tctcccacct 480
ccanagaggg gagaggggct 500

<210> 436
<211> 386
<212> DNA
<213> Homo sapiens

<400> 436
gtgctcatcc tgaactgtta ctccaaatcc actccgtttt taaagcaaaa ttatcttgtg 60
attttaagaa aagagttttc tatttattta agaaagtaac aatgcagtct gcaagctttc 120
agtagttttc tagtgctata ttcatcctgt aaaactctta ctacgtaacc agtaatcaca 180
aggaaagtgt ccccttgca tatttcttta aaattctttc tttggaaagt atgatgttga 240
taattaactt accttatct gccaaaacca gagcaaaatg ctaaatacgt tattgctaat 300
cagtggtctc aaatcgattt gcctcccttt gcctcgtctg agggctgtaa gcctgaagat 360
agtggcaagc accaagtcag tttcca 386

<210> 437
<211> 180
<212> DNA
<213> Homo sapiens

<400> 437
aaattgtctg tctcctatag cagaaaggtg aatgtacaaa ctgttggtgg ccctgaatcc 60
atctgaccag ctgctggtat ctgccaggac tggcagttct gatttagtta ggagagagcc 120
gctgataggt taggtctcat ttggagtgtt ggtggaaagg aaactgaagg taattgaata 180

<210> 438
<211> 570
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 11
<223> n = A,T,C or G

<400> 438
tcaagattta nccaaggctg tggcaaaggt gtaacttgta aacttgagtt ggagtactat 60
atttacaaat aaaattggca ccatgtgcca tctgtacata ttactgttgc atttactttt 120
aataaagctt gtggcccctt ttacttttt atagcttaac taatttgaat gtggttactt 180
cctactgtag ggtagcggaa aagttgtctt aaaaggtatg gtggggatat ttttaaaaac 240
tccttttggt ttacctgggg atccaattga tgtatatgtt tatatactgg gttcttgttt 300
tatatacctg gcttttactt tattaatatg agttactgaa ggtgatggag gtatttgaaa 360
attttacttc cataggacat actgcatgta agccaagtca tggagaatct gctgcatagc 420
tctattttaa agtaaaagtc taccaccgaa tccctagtcc ccctgttttc tgtttcttct 480
tgtgattgct gccataattc taagttattt acttttacca ctatttaagt tatcaacttt 540
agctagtatc ttcaaacttt cactttgaaa 570

<210> 439
<211> 551
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 11, 12
<223> n = A,T,C or G

<400> 439
ccaacacagt nntgaaacat tgttttgaat cctctgtaaa ccaaggcatt aatcttaata 60
aaccaggatc catttaggta ccacttgata taaaaaggat atccataatg aatattttat 120
actgcatcct ttacattagc cactaaatac gttattgctt gatgaagacc tttcacagaa 180
tcctatggat tgcagcattt cacttggcta cttcataccc atgccttaaa gaggggcagt 240
ttctcaaaag cagaaacatg ccgccagttc tcaagttttc ctcctaactc catttgaatg 300
taagggcagc tggcccccaa tgtggggagg tccgaacatt ttctgaattc ccattttctt 360
gttcgcggct aaatgacagt ttctgtcatt acttagattc cgatctttcc caaaggtgtt 420
gatttacaaa gaggccagct aatagcaaga aatcatgacc ctgaaagaga gatgaaattc 480
aagctgtgag ccaggcagga gctcagtatg gcaaaggttc ttgagaatca gccatttggt 540
acaaaaaaga t 551

<210> 440
<211> 464
<212> DNA
<213> Homo sapiens

<400> 440
cagtgtggtg gaattcaata actaaaaggt atgcaatcaa atctgctttt taaagaatgc 60
tctttacttc atggacttcc actgccatcc tcccaagggg cccaaattct ttcagtggct 120
acctacatac aattccaaac acatacagga aggtagaaat atctgaaaat gtatgtgtaa 180
gtattcttat ttaatgaaag actgtacaaa gtagaagtct tagatgtata tatttcctat 240
attgttttca gtgtacatgg aataacatgt aattaagtac tatgtatcaa tgagtaacag 300
gaaaatttta aaaatacaga tagatatatg ctctgcatgt tacataagat aaatgtgctg 360
aatggttttc aaaataaaaa tgaggtactc tcctggaaat attaagaaag actatctaaa 420
tgttgaaaga ccaaaaggtt aataaagtaa ttataactaa aaaa 464

<210> 441
<211> 485
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 243
<223> n = A,T,C or G

<400> 441
gattcactgg ggcattattt tgttagagga ccttaaaatt gtttatttttt taaatgtgat 60
tcctttatgg cattagggta aagatgaagc aataattttt aaattgtgta tgtgcatatg 120
aagcacagac atgcatgtgt gtgtgtgtct gtgtgtgtgt gtccgtgtat gtgtgtgtgg 180
gttctaatgg taatttgcct cagtcatttt tttaatattt gcagtacttg atttaggatc 240
tgnggcgcag ggcaatgttt caaagtttag tcacagctta aaaacattca gtgtgacttt 300
aatattataa aatgatttcc catgccataa tttttctgtc tattaaatgg gacaagtgta 360
aagcatgcaa aagttagaga tctgttatat aacatttgtt ttgtgatttg aactcctagg 420
aaaaatatga tttcataaat gtaaaatgca cagaaatgca tgcaatactt ataagactta 480
aaaat 485

<210> 442
<211> 334
<212> DNA
<213> Homo sapiens

<400> 442
ttgccagaat attccaagac atgttttaga agctacctat ggcattaaca tcataacgcc 60
tagagaggat gaagatcccc accgacctcc aacatcggaa gaactgttga cagcttatgg 120
atacatgcga ggattcatga cagcgcatgg acagccagac cagcctcgat ctgcgcgcta 180
catcctgaag gactatgtca gtggtaagct gctgtactgc catcctcctc ctggaagaga 240
tcctgtaact tttcagcatc aacaccagcg actcctagag aacaaaatga acagtgatga 300
aataaaaatg cagctaggca gaaataaaaa agca 334

<210> 443
<211> 235
<212> DNA
<213> Homo sapiens

<400> 443
atatgaaaat gtaaatatca cttgtgtact caaacaaaag ttggtcttaa gcttccacct 60
tgagcagcct tggaaaccta acctgcctct tttagcataa tcacattttc taaatgattt 120
tctttgttcc tgaaaaagtg atttgtatta gttttacatt tgtttttgg aagattatat 180
ttgtatatgt atcatcataa aatatttaaa taaaaagtat cttgagtgac aaaaa 235

<210> 444
<211> 297
<212> DNA
<213> Homo sapiens

<400> 444
taagtcaact gcttctgaaa taactctgta ttgtagatta tgcagatctt tacaggcata 60
aatatttaaa ctgtaatatg ctaacttgaa gagattgcaa taaagctgct tcagctaacc 120
ctgtttatgt ttaaatacta gggtttgttc tatattttat acatgcattt tggatgatta 180
aagaatgcct ggttttcgtt tgcaatttgc ttgtgtaaat caggttgtaa aaaggcagat 240
aaattgaaat gtttgtggta tgaggaaata aaagaatgga attagctttc aaaaaaa 297

<210> 445
<211> 344
<212> DNA
<213> Homo sapiens

<400> 445
gacttttgtt tagtgataga agatttgggg aggacccaaa ggactcagaa ctttctctcc 60
atacctcctt ttactctttt ctttctgtgt aatgtatcaa caactgttta atctcccttc 120
taacaaacct tgatataagc tttctgatat caaagtatat tgacagttaa cccttactga 180
ttttaaactt gactatccag tctgttaatt acctaagatt ttgttttcat ttcatctcta 240
attgttttga tcattggcag agaaagagta tttgaaattc atatcagttt tgctccttat 300
tttaatctct ttgaattaaa aataaaactt tttcaaaatg gaaa 344

<210> 446
<211> 294
<212> DNA
<213> Homo sapiens

<400> 446
tatcagatcc tttgaaaaga gaatatttac aatatatgac taatttgggg aaaatgaagt 60
tttgatttat ttgtgtttaa atgctgctgt cagacgattg ttcttagacc tcctaaatgc 120
cccatattaa aagaactcat tcataggaag gtgtttcatt ttggtgtgca accctgtcat 180

tacgtcaacg caacgtctaa ctggacttcc caagataaat ggtaccagcg tcctcttaaa 240
agatgcctta atccattcct tgaggacaga ccttagttga aatgatagca gaat 294

<210> 447
<211> 355
<212> DNA
<213> Homo sapiens

<400> 447
gcagtttgat ttaaaagtgt cactcttcct ccttttctac tttcagtaga tatgagatag 60
agcataatta tctgttttat cttagtttta tacataattt accatcagat agaactttat 120
ggttctagta cagatactct actacactca gcctcttatg tgccaagttt ttctttaagc 180
aatgagaaat tgctcatgtt cttcatcttc tcaaatcatc agaggccgaa gaaaaacact 240
ttggctgtgt ctataacttg acacagtcaa tagaatgaag aaaattagag tagttatgtg 300
attatttcag ctcttgacct gtcccctctg gctgcctctg agtctgaatc tccca 355

<210> 448
<211> 420
<212> DNA
<213> Homo sapiens

<400> 448
ccagtgtggt ggaattcgct tggaggttgg cggcgcgggg ctgaaggcta gcaaaccgag 60
cgatcatgtc gcacaaacaa atttactatt cggacaaata cgacgacgag gagtttgagt 120
atcgacatgt catgctgccc aaggacatag ccaagctggt ccctaaaacc catctgatgt 180
ctgaatctga atggaggaat cttggcgttc agcagagtca gggatgggtc cattatatga 240
tccatgaacc agaacctcac atcttgctgt tccggcgccc actacccaag aaaccaaaga 300
aatgaagctg gcaagctact tttcagcctc aagctttaca cagctgtcct tacttcctaa 360
catctttctg ataacattat tatgttgcct tcttgtttct cactttgata tttaaaagat 420

<210> 449
<211> 282
<212> DNA
<213> Homo sapiens

<400> 449
ccagtgtggt ggaattctgc agctcttggg ttttttgtgg cttccttcgt tattggagcc 60
aggcctacac cccagcaacc atgtccaagg gacctgcagt tggtattgat cttggcacca 120
cctactcttg tgtgggtgtt ttccagcacg gaaaagtcga gataattgcc aatgatcagg 180
gaaaccgaac cactccaagc tatgtcgcct ttacggacac tgaacggttg atcggtgatg 240
ccgcaaagaa tcaagttgca atgaacccca ccaacacagt tt 282

<210> 450
<211> 184
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 4, 11, 25, 33, 41, 43, 79, 86, 133, 147, 177, 182
<223> n = A,T,C or G

<400> 450
gcangatggc nctgatccaa atggnctcct tgnaggaggc ngnccacgcc ctcattgacc 60
tgcacaacca ctacctcgng gagaancacc acctgcgggt ctccttctcc aagtccacca 120
tctaggggca cangccccca cggacgntcc ccctggtgac aacttccatc attccanaga 180
anat 184

<210> 451
<211> 3188
<212> DNA
<213> Homo sapiens

<400> 451
attccggcgg ctccactccg tcccccgcgg tctgctctgt gtgccatgga cggcatcgtc 60
ccagatatag ccgttggtac aaagcgggga tctgacgagc ttttctctac ttgtgtcact 120
aacggaccgt ttatcatgag cagcaactcg gcttctgcag caaacggaaa tgacagcaag 180
aagttcaaag gtgacagccg aagtgcaggc gtcccctcta gagtgatcca catccggaag 240
ctccccatcg acgtcacgga gggggaagtc atctccctgg ggctgccctt tgggaaggtc 300
accaacctcc tgatgctgaa ggggaaaaac caggccttca tcgagatgaa cacggaggag 360
gctgccaaca ccatggtgaa ctactacacc tcggtgaccc ctgtgctgcg cggccagccc 420
atctacatcc agttctctaa ccacaaggag ctgaagaccg acagctctcc caaccaggcg 480
cgggcccagg cggccctgca ggcggtgaac tcggtccagt cggggaacct ggccttggct 540
gcctcggcgg cggccgtgga cgcagggatg gcgatggccg ggcagagccc tgtgctcagg 600
atcatcgtgg agaacctctt ctaccctgtg accctggatg tgctgcacca gattttctcc 660
aagttcggca cagtgttgaa gatcatcacc ttcaccaaga acaaccagtt ccaggccctg 720
ctgcagtatg cggaccccgt gagcgcccag cacgccaagc tgtcgctgga cgggcagaac 780
atctacaacg cctgctgcac gctgcgcatc gacttttcca agctcaccag cctcaacgtc 840
aagtacaaca atgacaagag ccgtgactac acacgcccag acctgccttc cggggacagc 900
cagccctcgc tggaccagac catggccgcg gccttcgcct ctccgtatgc aggagctggt 960
ttccctccca cctttgccat tcctcaagct gcaggccttt ccgttccgaa cgtccacggc 1020
gccctggccc ccctggccat cccctcggcg gcggcggcag ctgcggcggc aggtcggatc 1080
gccatcccgg gcctggcggg ggcaggaaat tctgtattgc tggtcagcaa cctcaaccca 1140
gagagagtca caccccaaag cctctttatt cttttcggcg tctacggtga cgtgcagcgc 1200
gtgaagatcc tgttcaataa gaaggagaac gccctagtgc agatggcgga cggcaaccag 1260
gcccagctgg ccatgagcca cctgaacggg cacaagctgc acgggaagcc gatccgcatc 1320
acgctctcga agcaccagaa cgtgcagctg ccccgcgagg gccaggagga ccagggcctg 1380
accaaggact acggcaactc acccctgcac cgcttcaaga agccgggctc caagaacttc 1440
cagaacatat tcccgccctc ggccactctg cacctctcca acatcccgcc ctcagtctcc 1500
gaggaggatc tcaaggtcct gttttccagc aatgggggcg tcgtcaaagg attcaagttc 1560
ttccagaagg accgcaagat ggcactgatc cagatgggct ccgtggagga ggcggtccag 1620
gccctcattg acctgcacaa ccacgacctc ggggagaacc accacctgcg ggtctccttc 1680
tccaagtcca ccatctaggg gcacaggccc ccacggccgg gccccctggc gacaacttcc 1740
atcattccag agaaaagcca ctttaaaaac agctgaagtg accttagcag accagagatt 1800
ttatttttt aaagagaaat cagtttacct gtttttaaaa aaattaaatc tagttcacct 1860
tgctcaccct gcggtgacag ggacagctca ggctcttggt gactgtggca gcgggagttc 1920
ccggccctcc acacccgggg ccagaccctc ggggccatgc cttggtgggg cctgtgtcgg 1980
gcgtggggcc tgcaggtggg cgccccgacc acgacttggc ttccttgtgc cttaaaaaac 2040
ctgcctttcc tgcagccaca cacccacccg gggtgtcctg gggacccaag gggtgggggg 2100
gtcacaccag agagaggcag ggggcctggc cggctcctgc aggatcatgc agctggggcg 2160
cggcggccgc gctgcgagca ccccaacccc agccctctaa tcaagtcacg tgattctccc 2220
ttcaccccgc ccccagggcc ttcccttcct tgcccccagg cgggctcccc gctgctccag 2280
ctgcggactg gtcgacataa tctctgtatt atatactttg cagttgcaga cgtctgtgcc 2340
tagcaatatt tccagttgac caaatattct aatctttttt catttatatg caaaagaaat 2400
agttttaagt aacttttat agcaagatga tacaatggta tgagtgtaat ctaaacttcc 2460
ttgtggtatt accttgtatg ctgttacttt tattttattc cttgtaatta agtcacaggc 2520
aggacccagt ttccagagag caggcggggc cgcccagtgg gtcaggcaca gggagccccg 2580
gtcctatctt agagcccctg agcttcaggg aagggcggcg tgtcgcgcct ctggcatcgc 2640
tccggttgcc ttacaccacg ccttcacctg cagtcgccta gaaaacttgc tctcaaactt 2700
cagggttttt tcttcttcaa atttggacca aagtctcatt tctgtgtttt gcctgcctct 2760
gatgctggga cccggaaagc gggcgctcct gtctttgtgc tctttctacc gcccccgcgt 2820
cctgtcccgg gggctctcct aggatcccct ttccgtaaaa gcgtgtaaca agggtgtaaa 2880
tatttataat tttttatacc tgttgtgaga cccgaggggc ggcggcgcgg ttttttatgg 2940
tgacacaaat gtatattttg ctaacagcaa ttccaggctc agtattgtga ccgcggagcc 3000
acaggggacc ccacgcacat tccgtgcctt acccgatggc ttgtgacgcg gagagaaccg 3060

attaaaaccg tttgagaaac tcctcccttg tctagccctg tgttcgctgt ggacgctgta 3120
gacacaggtt ggccagtctg tacctggact tcgaataaat cttctgtatc ctcaaaaaaa 3180
aaaaaaaa 3188

<210> 452
<211> 550
<212> PRT
<213> Homo sapiens

<400> 452
Met Asp Gly Ile Val Pro Asp Ile Ala Val Gly Thr Lys Arg Gly Ser
1 5 10 15
Asp Glu Leu Phe Ser Thr Cys Val Thr Asn Gly Pro Phe Ile Met Ser
20 25 30
Ser Asn Ser Ala Ser Ala Ala Asn Gly Asn Asp Ser Lys Lys Phe Lys
35 40 45
Gly Asp Ser Arg Ser Ala Gly Val Pro Ser Arg Val Ile His Ile Arg
50 55 60
Lys Leu Pro Ile Asp Val Thr Glu Gly Glu Val Ile Ser Leu Gly Leu
65 70 75 80
Pro Phe Gly Lys Val Thr Asn Leu Leu Met Leu Lys Gly Lys Asn Gln
85 90 95
Ala Phe Ile Glu Met Asn Thr Glu Glu Ala Ala Asn Thr Met Val Asn
100 105 110
Tyr Tyr Thr Ser Val Thr Pro Val Leu Arg Gly Gln Pro Ile Tyr Ile
115 120 125
Gln Phe Ser Asn His Lys Glu Leu Lys Thr Asp Ser Ser Pro Asn Gln
130 135 140
Ala Arg Ala Gln Ala Ala Leu Gln Ala Val Asn Ser Val Gln Ser Gly
145 150 155 160
Asn Leu Ala Leu Ala Ala Ser Ala Ala Ala Val Asp Ala Gly Met Ala
165 170 175
Met Ala Gly Gln Ser Pro Val Leu Arg Ile Ile Val Glu Asn Leu Phe
180 185 190
Tyr Pro Val Thr Leu Asp Val Leu His Gln Ile Phe Ser Lys Phe Gly
195 200 205
Thr Val Leu Lys Ile Ile Thr Phe Thr Lys Asn Asn Gln Phe Gln Ala
210 215 220
Leu Leu Gln Tyr Ala Asp Pro Val Ser Ala Gln His Ala Lys Leu Ser
225 230 235 240
Leu Asp Gly Gln Asn Ile Tyr Asn Ala Cys Cys Thr Leu Arg Ile Asp
245 250 255
Phe Ser Lys Leu Thr Ser Leu Asn Val Lys Tyr Asn Asn Asp Lys Ser
260 265 270
Arg Asp Tyr Thr Arg Pro Asp Leu Pro Ser Gly Asp Ser Gln Pro Ser
275 280 285
Leu Asp Gln Thr Met Ala Ala Ala Phe Ala Ser Pro Tyr Ala Gly Ala
290 295 300
Gly Phe Pro Pro Thr Phe Ala Ile Pro Gln Ala Ala Gly Leu Ser Val
305 310 315 320
Pro Asn Val His Gly Ala Leu Ala Pro Leu Ala Ile Pro Ser Ala Ala
325 330 335
Ala Ala Ala Ala Ala Ala Gly Arg Ile Ala Ile Pro Gly Leu Ala Gly
340 345 350
Ala Gly Asn Ser Val Leu Leu Val Ser Asn Leu Asn Pro Glu Arg Val
355 360 365
Thr Pro Gln Ser Leu Phe Ile Leu Phe Gly Val Tyr Gly Asp Val Gln
370 375 380

```
Arg Val Lys Ile Leu Phe Asn Lys Lys Glu Asn Ala Leu Val Gln Met
385                 390                 395                 400
Ala Asp Gly Asn Gln Ala Gln Leu Ala Met Ser His Leu Asn Gly His
                405                 410                 415
Lys Leu His Gly Lys Pro Ile Arg Ile Thr Leu Ser Lys His Gln Asn
            420                 425                 430
Val Gln Leu Pro Arg Glu Gly Gln Glu Asp Gln Gly Leu Thr Lys Asp
        435                 440                 445
Tyr Gly Asn Ser Pro Leu His Arg Phe Lys Lys Pro Gly Ser Lys Asn
    450                 455                 460
Phe Gln Asn Ile Phe Pro Pro Ser Ala Thr Leu His Leu Ser Asn Ile
465                 470                 475                 480
Pro Pro Ser Val Ser Glu Glu Asp Leu Lys Val Leu Phe Ser Ser Asn
                485                 490                 495
Gly Gly Val Val Lys Gly Phe Lys Phe Phe Gln Lys Asp Arg Lys Met
            500                 505                 510
Ala Leu Ile Gln Met Gly Ser Val Glu Glu Ala Val Gln Ala Leu Ile
        515                 520                 525
Asp Leu His Asn His Asp Leu Gly Glu Asn His His Leu Arg Val Ser
    530                 535                 540
Phe Ser Lys Ser Thr Ile
545                 550
```

<210> 453
<211> 2257
<212> DNA
<213> Homo sapiens

<400> 453

```
attttgctta cagagtcccg tctcaccatc ctgggcttcc aacggagact gcggtatccg 60
cggctggaga cccagcggcg agtagccttt tgctcccgga cggacttgag aggcttaaag 120
gatggcctcg tcagatctgg aacaattatg ctctcatgtt aatgaaaaga ttggcaatat 180
taagaaaacc ttatcattaa gaaactgtgg ccaggaacct accttgaaaa ctgtattaaa 240
taaaatagga gatgagatca ttgtaataaa tgaacttcta aataaattgg aattggaaat 300
tcagtatcaa gaacaaacca acaattcact caaggaactc tgtgaatctc ttgaagaaga 360
ttacaaagac atagaacatc ttaaagaaaa cgttccttcc catttgcctc aagtaacagt 420
aacccagagc tgtgttaagg gatcagatct tgatcctgaa gaaccaatca aagttgaaga 480
acctgaaccc gtaaagaagc ctcccaaaga gcaaagaagt attaaggaaa tgccatttat 540
aacttgtgat gagttcaatg gtgttccttc gtacatgaaa tcccgcttaa cctataatca 600
aattaatgat gttattaaag aaatcaacaa ggcagtaatt agtaaatata aaatcctaca 660
tcagccaaaa aagtctatga attctgtgac cagaaatctc tatcacagat ttattgatga 720
agaaacgaag gataccaaag gtcgttattt tatagtggaa gctgacataa aggagttcac 780
aactttgaaa gctgacaaga agtttcacgt gttactgaat attttacgac actgccggag 840
gctatcagag gtccgagggg gaggacttac tcgttatgtt ataacctgag tcccttgtga 900
acttttgaac ataccaacag ggtatagagt atagaggcta tttctataat tttcttatat 960
ataatttttt taacttttaa tctttttttgt ttcctttttt ttttttttga gacaggatct 1020
tgctttgtca cccaggggct tgctttgtca cgcaggctag agtgcagtgg cgcaaacatg 1080
gctcactgca gcctcaacct cccaggctca agtgatcctc ccacctcagc cccctgaatg 1140
gctgggacta caagcgtgcg ccaccatgcc tggctaattt ttgtattttt tggagagatg 1200
gggtttcacc atgttgccta ggctggtctt gagctcctga gctcaaacaa tccaccctcc 1260
tcagcctccc aaagtgctgg gattacaggc ttgagccacc acacctgacc tattcttgtt 1320
tcttataaaa ataaaacttt tttggataaa gcttatttct tgtttttttc tttttctttt 1380
ttttttttt tcgagactcc atctcagaaa aaaagaaaaa aagactgggt acagatgtga 1440
tattggaaga aaaagatcaa gctgatgagg ttaggatacc caggcccttt ggacttaaag 1500
atcactagtg tctaaattcc atcgatggca tttcagtcta taggtaaact tcctggaagc 1560
tggatttgga gacagtttat catctgatta ttgggctttc gtataggtcc ttagggagca 1620
gcttacctga aatgcattta gtgtacacca gtctgtaaac ttcaacctgt aatgaaagtg 1680
```

taataaatgt acattgagtt gatgtgataa tgtgatataa taagaaatat atatttgatc 1740
ttcctatcta gttccttgtt cagagctcct aaaaccttg taatttccaa agtgatggag 1800
tacatctttt gttctagtat ttggtctttg accccagttc ctgacacaaa gctcctaaat 1860
tcctttaaat ttcccagtga taggagaatt ttttgttcta atgaggtcac tcttgatggg 1920
cacctggata actcaggatg ggggctgctc acaaagacca catcatgatt ggaagtttca 1980
aactttcagt ctcccacctc cagagagggg agaggggctg gagatttgtg tcaataatcc 2040
atcaggccta tgtcaacaag acataatccg ttaactatgg agttcaggga gcttcagggt 2100
tggcaaacat tttgatgtgc caggaaggtg acgcactcca gctttatgaa gtcagcaagt 2160
cctgtgctca ggatgcttyt ggaccttgcc ccaggtaccc cttcatgtgg ctgttgttca 2220
tctgtatcct ttgtagtagc cttaaaataa actgtta 2257

<210> 454
<211> 255
<212> PRT
<213> Homo sapiens

<400> 454
Met Ala Ser Ser Asp Leu Glu Gln Leu Cys Ser His Val Asn Glu Lys
1 5 10 15
Ile Gly Asn Ile Lys Lys Thr Leu Ser Leu Arg Asn Cys Gly Gln Glu
20 25 30
Pro Thr Leu Lys Thr Val Leu Asn Lys Ile Gly Asp Glu Ile Ile Val
35 40 45
Ile Asn Glu Leu Leu Asn Lys Leu Glu Leu Glu Ile Gln Tyr Gln Glu
50 55 60
Gln Thr Asn Asn Ser Leu Lys Glu Leu Cys Glu Ser Leu Glu Glu Asp
65 70 75 80
Tyr Lys Asp Ile Glu His Leu Lys Glu Asn Val Pro Ser His Leu Pro
85 90 95
Gln Val Thr Val Thr Gln Ser Cys Val Lys Gly Ser Asp Leu Asp Pro
100 105 110
Glu Glu Pro Ile Lys Val Glu Glu Pro Glu Pro Val Lys Lys Pro Pro
115 120 125
Lys Glu Gln Arg Ser Ile Lys Glu Met Pro Phe Ile Thr Cys Asp Glu
130 135 140
Phe Asn Gly Val Pro Ser Tyr Met Lys Ser Arg Leu Thr Tyr Asn Gln
145 150 155 160
Ile Asn Asp Val Ile Lys Glu Ile Asn Lys Ala Val Ile Ser Lys Tyr
165 170 175
Lys Ile Leu His Gln Pro Lys Lys Ser Met Asn Ser Val Thr Arg Asn
180 185 190
Leu Tyr His Arg Phe Ile Asp Glu Glu Thr Lys Asp Thr Lys Gly Arg
195 200 205
Tyr Phe Ile Val Glu Ala Asp Ile Lys Glu Phe Thr Thr Leu Lys Ala
210 215 220
Asp Lys Lys Phe His Val Leu Leu Asn Ile Leu Arg His Cys Arg Arg
225 230 235 240
Leu Ser Glu Val Arg Gly Gly Gly Leu Thr Arg Tyr Val Ile Thr
245 250 255

<210> 455
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 455
gcctcgtcag atctggaaca attatgctc 29

<210> 456
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 456
cgtaactcga gtcatcaggt tataacataa c 31

<210> 457
<211> 262
<212> PRT
<213> Homo sapiens

<400> 457
Met Gln His His His His His His Ala Ser Ser Asp Leu Glu Gln Leu
1 5 10 15
Cys Ser His Val Asn Glu Lys Ile Gly Asn Ile Lys Lys Thr Leu Ser
20 25 30
Leu Arg Asn Cys Gly Gln Glu Pro Thr Leu Lys Thr Val Leu Asn Lys
35 40 45
Ile Gly Asp Glu Ile Ile Val Ile Asn Glu Leu Leu Asn Lys Leu Glu
50 55 60
Leu Glu Ile Gln Tyr Gln Glu Gln Thr Asn Asn Ser Leu Lys Glu Leu
65 70 75 80
Cys Glu Ser Leu Glu Glu Asp Tyr Lys Asp Ile Glu His Leu Lys Glu
85 90 95
Asn Val Pro Ser His Leu Pro Gln Val Thr Val Thr Gln Ser Cys Val
100 105 110
Lys Gly Ser Asp Leu Asp Pro Glu Glu Pro Ile Lys Val Glu Glu Pro
115 120 125
Glu Pro Val Lys Lys Pro Pro Lys Glu Gln Arg Ser Ile Lys Glu Met
130 135 140
Pro Phe Ile Thr Cys Asp Glu Phe Asn Gly Val Pro Ser Tyr Met Lys
145 150 155 160
Ser Arg Leu Thr Tyr Asn Gln Ile Asn Asp Val Ile Lys Glu Ile Asn
165 170 175
Lys Ala Val Ile Ser Lys Tyr Lys Ile Leu His Gln Pro Lys Lys Ser
180 185 190
Met Asn Ser Val Thr Arg Asn Leu Tyr His Arg Phe Ile Asp Glu Glu
195 200 205
Thr Lys Asp Thr Lys Gly Arg Tyr Phe Ile Val Glu Ala Asp Ile Lys
210 215 220
Glu Phe Thr Thr Leu Lys Ala Asp Lys Lys Phe His Val Leu Leu Asn
225 230 235 240
Ile Leu Arg His Cys Arg Arg Leu Ser Glu Val Arg Gly Gly Gly Leu
245 250 255
Thr Arg Tyr Val Ile Thr
260

<210> 458

<211> 792
<212> DNA
<213> Homo sapiens

<400> 458
atgcagcatc accaccatca ccacgcctcg tcagatctgg aacaattatg ctctcatgtt 60
aatgaaaaga ttggcaatat taagaaaacc ttatcattaa gaaactgtgg ccaggaacct 120
accttgaaaa ctgtattaaa taaaatagga gatgagatca ttgtaataaa tgaacttcta 180
aataaattgg aattggaaat tcagtatcaa gaacaaacca acaattcact caaggaactc 240
tgtgaatctc ttgaagaaga ttacaaagac atagaacatc ttaaagaaaa cgttccttcc 300
catttgcctc aagtaacagt aacccagagc tgtgttaagg gatcagatct tgatcctgaa 360
gaaccaatca aagttgaaga acctgaaccc gtaaagaagc ctcccaaaga gcaaagaagt 420
attaaggaaa tgccatttat aacttgtgat gagttcaatg gtgttccttc gtacatgaaa 480
tcccgcttaa cctataatca aattaatgat gttattaaag aaatcaacaa ggcagtaatt 540
agtaaatata aaatcctaca tcagccaaaa aagtctatga attctgtgac cagaaatctc 600
tatcacagat ttattgatga agaaacgaag gataccaaag gtcgttattt tatagtggaa 660
gctgacataa aggagttcac aactttgaaa gctgacaaga agtttcacgt gttactgaat 720
attttacgac actgccggag gctatcagag gtccgagggg gaggacttac tcgttatgtt 780
ataacctgat ga 792

<210> 459
<211> 15
<212> PRT
<213> Homo sapiens

<400> 459
Lys Glu Leu Cys Glu Ser Leu Glu Glu Asp Tyr Lys Asp Ile Glu
1 5 10 15

<210> 460
<211> 15
<212> PRT
<213> Homo sapiens

<400> 460
Asp Pro Glu Glu Pro Ile Lys Val Glu Glu Pro Glu Pro Val Lys
1 5 10 15

<210> 461
<211> 15
<212> PRT
<213> Homo sapiens

<400> 461
Met Ala Ser Ser Asp Leu Glu Gln Leu Cys Ser His Val Asn Glu
1 5 10 15

<210> 462
<211> 15
<212> PRT
<213> Homo sapiens

<400> 462
Lys Ile Gly Asp Glu Ile Ile Val Ile Asn Glu Leu Leu Asn Lys
1 5 10 15

<210> 463
<211> 15
<212> PRT
<213> Homo sapiens

<400> 463
Thr Leu Lys Ala Asp Lys Lys Phe His Val Leu Leu Asn Ile Leu
1 5 10 15

<210> 464
<211> 20
<212> PRT
<213> Homo sapiens

<400> 464
Ala Val Ile Ser Lys Tyr Lys Ile Leu His Gln Pro Lys Lys Ser Met
1 5 10 15
Asn Ser Val Thr
20

<210> 465
<211> 20
<212> PRT
<213> Homo sapiens

<400> 465
Leu Thr Tyr Asn Gln Ile Asn Asp Val Ile Lys Glu Ile Asn Lys Ala
1 5 10 15
Val Ile Ser Lys
20

<210> 466
<211> 20
<212> PRT
<213> Homo sapiens

<400> 466
Ile Asn Asp Val Ile Lys Glu Ile Asn Lys Ala Val Ile Ser Lys Tyr
1 5 10 15
Lys Ile Leu His
20

<210> 467
<211> 20
<212> PRT
<213> Homo sapiens

<400> 467
Lys Glu Ile Asn Lys Ala Val Ile Ser Lys Tyr Lys Ile Leu His Gln
1 5 10 15
Pro Lys Lys Ser
20

<210> 468
<211> 20
<212> PRT
<213> Homo sapiens

<400> 468
Tyr Met Lys Ser Arg Leu Thr Tyr Asn Gln Ile Asn Asp Val Ile Lys
1 5 10 15
Glu Ile Asn Lys
20

<210> 469
<211> 20
<212> PRT
<213> Homo sapiens

<400> 469
Asn Gly Val Pro Ser Tyr Met Lys Ser Arg Leu Thr Tyr Asn Gln Ile
1 5 10 15
Asn Asp Val Ile
20

<210> 470
<211> 20
<212> PRT
<213> Homo sapiens

<400> 470
Lys Ile Gly Asp Glu Ile Ile Val Ile Asn Glu Leu Leu Asn Lys Leu
1 5 10 15
Glu Leu Glu Ile
20

<210> 471
<211> 20
<212> PRT
<213> Homo sapiens

<400> 471
Lys Thr Val Leu Asn Lys Ile Gly Asp Glu Ile Ile Val Ile Asn Glu
1 5 10 15
Leu Leu Asn Lys
20

<210> 472
<211> 20
<212> PRT
<213> Homo sapiens

<400> 472
Lys Ile Gly Asn Ile Lys Lys Thr Leu Ser Leu Arg Asn Cys Gly Gln
1 5 10 15

Glu Pro Thr Leu
20

<210> 473
<211> 20
<212> PRT
<213> Homo sapiens

<400> 473
Ser His Val Asn Glu Lys Ile Gly Asn Ile Lys Lys Thr Leu Ser Leu
1 5 10 15
Arg Asn Cys Gly
20